

UNIVERSIDAD CARLOS III DE MADRID

**INGENIERIA TÉCNICA TELECOMUNICACIONES
ESPECIALIDAD TELEMÁTICA**

PROYECTO FIN DE CARRERA

Desarrollo de un Plugin de Firefox para el análisis de la publicidad online
basada en comportamiento



Autor: Humberto Hernández Ansorena

Tutor: Rubén Cuevas Rumín

Octubre 2015

Agradecimientos

En primer lugar quiero agradecer a mis padres por todo su apoyo y ánimo incondicional durante toda mi vida, y a mi esposa Emma y mi hijo David, ellos son los que me han dado las fuerzas necesarias para afrontar este reto en esta nueva etapa de mi vida donde el tiempo es más limitado y las ganas no son las mismas. Os quiero.

También querría agradecerle a mi tutor Rubén su ayuda en el desarrollo del proyecto.

ÍNDICE DE CONTENIDO

1. Introducción.....	4
1.1. Objetivos.....	6
2. Herramientas relacionadas.....	7
2.1. Herramientas basadas en extensiones.....	7
2.2. Herramientas basadas en proxy.....	9
3. Material y metodología.....	11
3.1. Herramientas y tecnologías utilizadas en el servidor.....	11
3.1.1. XAMMP.....	11
3.1.2. Apache	12
3.1.3. MySQL.....	12
3.1.4. PHP.....	15
3.1.5. Eclipse	16
3.1.6. Java.....	18
3.2. Herramientas y tecnologías utilizadas en el cliente.....	19
3.2.1. Firefox.....	19
3.2.2. Ajax.....	19
3.2.3. Javascript.....	21
3.2.4. Extensiones.....	23
4. Resultados.....	25
4.1. Casos de uso.....	25
4.2. Batería de pruebas.....	25
5. Planificación y presupuesto.....	29
5.1. Fases del proyecto.....	29
5.2. Personal.....	30
5.3. Presupuesto.....	31
5.3.1. Estimación de tiempos.....	31

5.3.2. Estimación de costes.....	31
6. Conclusiones y futuros desarrollos.....	32
6.1. Conclusiones.....	32
6.2. Futuros desarrollos.....	32
6.3. Dificultades encontradas.....	33
7. Referencias.....	36
Anexo 1. Manual del desarrollador.....	39
Anexo 2. Instalación del SDK Add-on.....	47
Anexo 3. Creación de un Add-on.....	49
Anexo 4. Librerías utilizadas en el desarrollo del Add-on.....	51

1. INTRODUCCIÓN

Un estudio reciente de la IAB (Interactive Advertising Bureau) reveló que la publicidad online en 2013 generó unos ingresos de 42 billones de dólares, un incremento del 17% con respecto al año anterior. Este fuerte crecimiento hace que la publicidad hoy en día sea una de las principales fuentes de ingreso de los servicios web. Algunos de los principales servicios tales como buscadores, redes sociales y páginas web incluyen publicidad en sus páginas.



Una de las ventajas que ofrece la publicidad online con respecto a la publicidad tradicional, que se realiza a través de la TV, es la capacidad de analizar de manera individual a cada uno de los usuarios web y ofrecerle la publicidad según sus intereses específicos. En los últimos años se han desarrollado una serie de técnicas en las cuales se recoge la información de navegación referente al usuario (historial de navegación, cookies, etc). Dicha información es analizada y utilizada por las empresas de publicidad para enviar anuncios.

La falta de transparencia, en cuanto a las técnicas empleadas y el tipo de información obtenida por las compañías sobre los usuarios, está creando preocupación en la sociedad debido a que se está invadiendo la privacidad y obteniendo información sensible de los usuarios. Actualmente se han creado una serie de iniciativas (W3C's Do-no-Track, Mozilla's Cookie Clearing House,...) a favor de la defensa de la privacidad, algunas de las cuales han ido más allá y han creado una serie de herramientas software que impiden cualquier transferencia de información de los usuarios finales hacia el ecosistema de la publicidad online. Esta falta de transparencia por parte de las compañías de publicidad puede provocar que los usuarios finales pierdan su confianza y reaccionen negativamente, interrumpiendo así la economía digital y afectando a su vez al propio sector de la publicidad online.

Esto tendría consecuencias muy negativas sobre la innovación en los servicios y aplicaciones en internet, una interrupción de modelos de freemium y, en casos extremos, la desaparición de ciertos servicios de internet que serían incapaces de sobrevivir sin los ingresos de la publicidad.

A su vez estos efectos tendrían un impacto social negativo representado por la pérdida de puestos de trabajo directos e indirectos vinculados a la industria de la publicidad online, además de privar a los consumidores europeos de los contenidos y servicios online con publicidad.

Para poder evitar todo lo mencionado con anterioridad y mantener un crecimiento sostenible, el mercado de la publicidad online debe evolucionar para integrar la transparencia y la preservación de la privacidad como eje fundamental para mantener la confianza de los usuarios. Para ello, el objetivo principal del Proyecto Europeo TYPES (Towards transparency and privacy in the online advertasing business) es definir un marco integral de tecnologías y herramientas que garanticen la transparencia y preservación de la privacidad, dándole el control a los propios usuarios finales para que sean ellos los que decidan la información que desean compartir.

1.1 Objetivo:

El objetivo del proyecto es la creación de una extensión para Firefox que nos permita recoger toda la información referente a la navegación del usuario, para posteriormente utilizarla en un estudio a nivel europeo. Dicho estudio determinará si se está vulnerando la privacidad de los usuarios a través de la publicidad, centrándose principalmente en toda aquella información de carácter sensible como pueden ser temas relacionados con la salud del usuario.

La información recogida se analizará posteriormente para tomar medidas legales contra todas aquellas compañías publicitarias que no cumplan con las normativas de la Unión Europea en relación a la privacidad de los usuarios de internet.



2. HERRAMIENTAS RELACIONADAS

Actualmente existen varias herramientas que permiten al usuario controlar y saber en todo momento quien esta accediendo a su información cuando esta navegando por internet. A continuación se mostraran algunas de las herramientas, las cuales se dividen en dos grupos: herramientas basadas en extensiones y herramientas basadas en proxy.

2.1 Herramientas basadas en extensiones (Browser add-on and extensions)

LIGHTBEAM



Es un complemento para Firefox que, a través de visualizaciones interactivas, muestra los sitios de origen y de terceros que interactúan contigo en la web. Lightbeam te muestra la web actual en toda su dimensión incluyendo aquellas partes no tan claras para el usuario medio. Dicha extensión crea un registro de eventos para cada sitio que visites y cada sitio de terceros que esté almacenado localmente en tu navegador. También nos ofrece la posibilidad de bloquear sitios específicos que nos estén rastreando, se toma nota de todas las peticiones que coincidan con una lista de servicios de seguimiento web y anunciantes.

GHOSTERY



Este complemento detecta rastreadores, balizas web, píxeles especiales e indicadores incluidos en páginas web por Facebook, Google y más de 500 servicios de todo tipo, como redes publicitarias, agencia de datos de comportamiento, editores de contenidos y todas aquellas empresas interesadas en la actividad de los usuarios.

Ghostery nos ofrece también la posibilidad de obtener más información acerca de cada una de las empresas que están analizando nuestra navegación, identificando a

las mismas e incluyendo enlaces a sus políticas de privacidad. Asimismo nos permite bloquear scripts, imágenes, objetos y documentos insertados en macros procedentes de empresas en las que no confíe.

GhostRank es una función exclusivamente opcional de Ghostery, la cual envía a sus servidores información anónima sobre los rastreadores que encuentre y su ubicación, permitiendo crear una lista más exhaustiva de elementos detectables. Así contribuimos a crear un ecosistema de publicidad basada en un comportamiento más transparente.

ADBLOCK PLUS



AdBlock Plus bloquea todo tipo de anuncios molestos en la web por defecto, incluyendo los anuncios de vídeo en youtube, los anuncios de Facebook, banners llamativos, pop-ups, pop-unders y mucho más.

Una gran cantidad de sitios web se basan en los ingresos por publicidad. Adblock Plus fomenta el uso de la publicidad no instructiva en lugar de anuncios irrelevantes molestos. Para lograr esto ha establecido una serie de estrictas normas conocidas como “Anuncios Aceptables”. Junto con nuestros usuarios y terceros contribuyentes hemos definido anuncios aceptables, permitiendo así que los anuncios que cumplan una serie de criterios sean visibles; se puede sin embargo desactivar esta opción en cualquier momento.

Otra de las funcionalidades que ofrece esta herramienta es la desactivación del seguimiento que realizan las empresas publicitarias de la navegación del usuario a través de las páginas que ha visitado y su historial de navegación.

PRIVACY BADGER



Privacy Badger analiza los sitios web visitados para detectar si estamos siendo analizados sin nuestro consentimiento. Esta herramienta toma nota de los dominios de terceros que incluyan imágenes, scripts y publicidad en las páginas que visitamos.

Si en un servidor de terceros se detecta que se está realizando un seguimiento sin permiso, mediante el uso de las cookies, Privacy Badger rechazará automáticamente el contenido de ese servidor. En algunos casos puede que un dominio de terceros nos ofrezca algún aspecto importante de la funcionalidad de la página como mapas incorporados, imágenes, etc. En esos casos Privacy Badger permitirá conexiones con terceros, pero nos mostrará sus cookies de seguimiento.

2.2 Herramientas basadas en proxy (Proxy-based solutions)

PRIVOX



Privoxy es un proxy web con capacidades avanzadas de filtrado para la mejora de la privacidad de los usuarios finales, modificando los datos de las páginas web y las cabeceras http, control de acceso y eliminación de todo tipo de anuncios desagradables que nos llegan a través de Internet.

Privoxy nos ofrece una configuración flexible y puede ser personalizado para adaptarse a las necesidades y gustos individuales. Se puede emplear tanto en sistemas autónomos como en redes multi-usuario.

AWAZZA



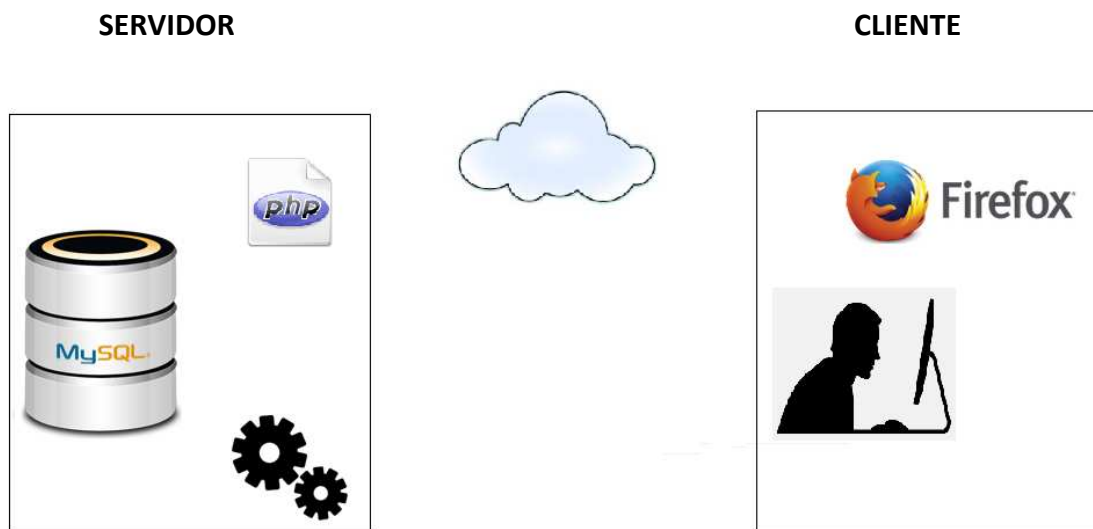
Awazza es el nuevo acelerador de navegación móvil de Telefónica I+D, quienes han logrado que mediante un cambio de proxy se aumente la velocidad 3G hasta en un 90 %, además de otros factores como priorizar los contenidos sobre la publicidad a la hora de cargar la página web.

Lo mejor es que Awazza funciona en cualquier plataforma, y con cualquier operador de España, no se limita únicamente a los clientes de Movistar. El "cómo" también es sencillo: funciona mediante varios servidores proxy que registran las webs más visitadas por sus usuarios para poder tenerlas precargadas y mostrarlas de forma casi instantánea. Para determinadas webs será una ventaja enorme, aunque para otras que se actualicen con mucha frecuencia, como un medio digital con contenido muy cambiante, no se podrá alcanzar una velocidad tan alta. Otro buen detalle sobre su funcionamiento: ha sido creado con la capacidad de cargar en último lugar la publicidad de cada sitio web, priorizando los contenidos.

3. MATERIAL Y METODOLOGÍA

El objetivo de este capítulo es el de analizar las herramientas y tecnologías que se han empleado en el desarrollo del proyecto. Dichas herramientas se dividirán en dos apartados: Herramientas utilizadas en el servidor y herramientas utilizadas en el cliente.

Diseño funcional de la aplicación



3.1 Herramientas y tecnologías utilizadas en el servidor

A continuación se hablará de todas las herramientas y tecnologías utilizadas en el lado del servidor.

3.1.1 XAMPP

XAMPP es una distribución de Apache completamente gratuita y fácil de instalar que contiene MySQL, PHP y Perl. El paquete de instalación de XAMPP ha sido diseñado para ser increíblemente fácil de instalar y usar. Solamente requiere descargar y ejecutar un archivo, con unas pocas configuraciones de algunos de sus componentes que el servidor web necesitara. XAMPP se actualiza regularmente para incorporar las últimas versiones de Apache/MySQL/PHP y Perl.

Oficialmente, los diseñadores de XAMPP, fueron los Baiker y Anthony Cooperation los cuales solo pretendían su uso como una herramienta de desarrollo, para permitir a los diseñadores de sitios webs y programadores testear su trabajo en sus propios ordenadores sin ningún acceso a Internet

3.1.2 APACHE

El servidor HTTP Apache es un servidor web HTTP de código abierto para múltiples plataformas que implementa el protocolo HTTP/1.1 y la noción de sitio virtual, se usa principalmente para enviar páginas web estáticas y dinámicas en la World Wide Web.

3.1.3 MYSQL

MySQL es uno de los sistemas de gestión de base de datos relacional más utilizados, cuenta con millones de aplicaciones. Es desarrollado, distribuido y mantenido por MySQL AB, compañía fundada por desarrolladores de MySQL, en enero de 2008 pasa a formar parte de Sun Microsystems y ésta a su vez de Oracle Corporation desde Abril de 2009.

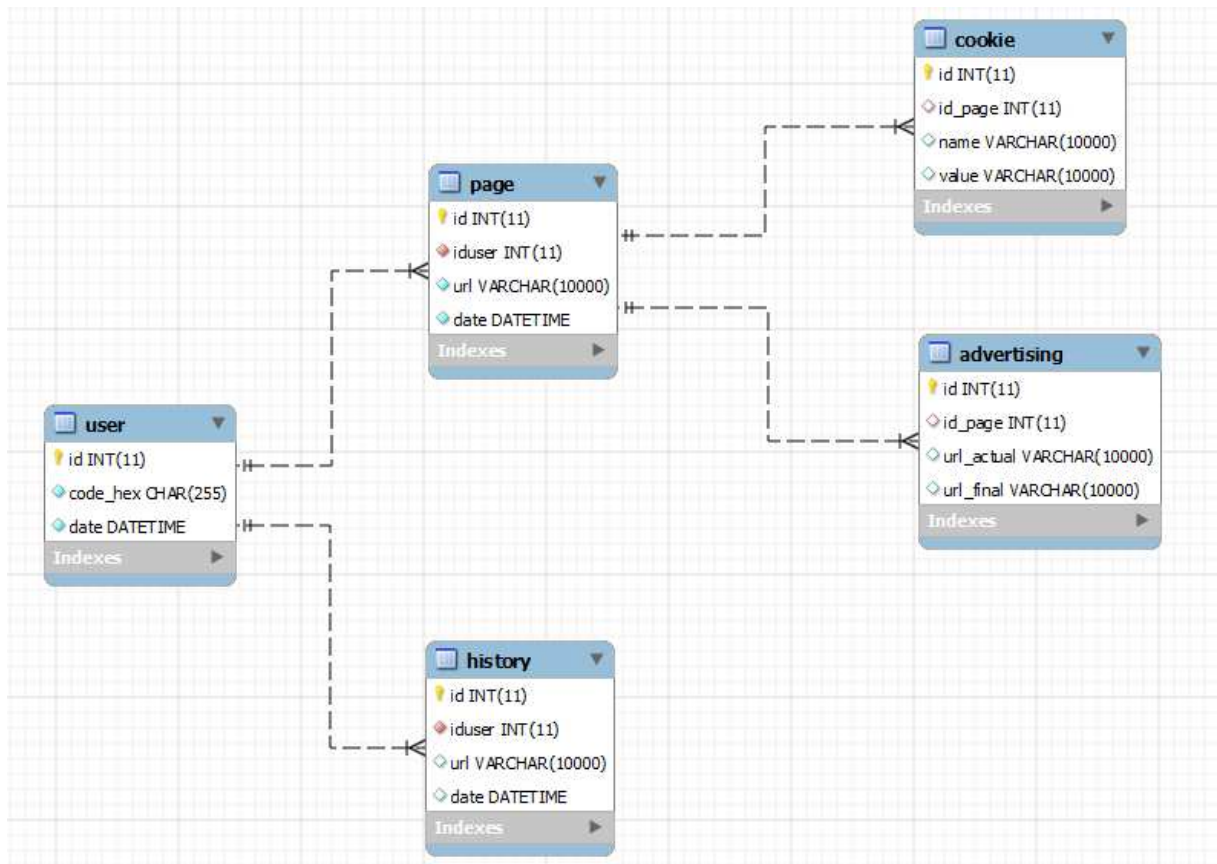
Oracle desarrolla MySQL como software libre bajo la licencia GNU PL, aquellas empresas que quieran incorporarlo en sus productos privativos deberán comprar a la empresa una licencia específica que les permita este uso, a diferencia de Apache donde el software es desarrollado por una comunidad pública. MySQL es patrocinado por una empresa privada la cual posee los derechos de la mayor parte del código.

MySQL se utiliza en sitios webs muy populares tales como Wikipedia, Google (no para realizar búsquedas), Facebook, Twitter, etc. Es una base de datos muy rápida en la lectura de información cuando utiliza el motor no transaccional MyISAM, sin embargo puede provocar problemas de integridad en entornos de alta concurrencia en la modificación, por esta razón se utiliza mucho en las aplicaciones web donde la lectura de los datos para mostrar la información de la página es más frecuente a diferencia de la modificación, la cual se hace de manera más puntual. Su popularidad como aplicación web está muy ligada a PHP que a menudo aparece en combinación con MySQL.

Otras de sus características más significativas y por la cual se elige MySQL es la por la gran cantidad de interfaces de programación que nos permiten acceder a la información de la base de datos (PHP, Python, Perl, Ruby, C, C++, C#, Java, etc), cada una de ellos utiliza una interfaz específica.

La incorporación de MySQL en nuestro proyecto se realiza a través de servidor XAMPP, el cual ya tiene instalado internamente el motor de base de datos. Una vez instalado el XAMPP con su motor de base de datos MySQL ya podemos crear las tablas en la base de datos y realizar las correspondientes operaciones de consulta, inserción, actualización y eliminación.

A continuación se presentará un esquema de la base de datos del proyecto, para la realización de dicho esquema se ha empleado la herramienta MySQL Workbench.



Veamos ahora de manera más detallada cada una de las tablas de la base de datos empleadas en el proyecto.

- **Tabla de usuarios:**

En esta tabla se almacenan todos y cada uno de los usuarios que están utilizando el plugin para navegar a través de Firefox, consta de un código hexadecimal que identifica de manera única al usuario y un campo fecha que refleja la fecha en la cual se ha dado de alta el usuario en la tabla.

#	Nombre	Tipo de datos	Longitud/Conjunto	Sin signo	Permitir NULL	Rellenar con ceros	Predeterminado
1	id	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREMENT
2	code_hex	CHAR	255	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Sin valor predeter...
3	date	DATETIME		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Sin valor predeter...

- **Tabla de páginas:**

Esta tabla recoge todas las páginas visitadas por el usuario a través del navegador Firefox, se almacena información referente a la URL de la página visitada así como la fecha en la cual se realizó la visita.

#	Nombre	Tipo de datos	Longitud/Conjunto	Sin signo	Permitir NULL	Rellenar con ceros	Predeterminado
1	id	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREMENT
2	iduser	INT	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sin valor predeter...
3	url	VARCHAR	10000	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Sin valor predeter...
4	date	DATETIME		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Sin valor predeter...

- **Tabla de cookies:**

Aquí se recopilan las cookies de todas y cada una de las páginas visitadas por el usuario, la información que se obtiene en relación a las cookies son: el nombre de la cookie y su valor correspondiente.

#	Nombre	Tipo de datos	Longitud/Conjunto	Sin signo	Permitir NULL	Rellenar con ceros	Predeterminado
1	id	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREMENT
2	id_page	INT	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sin valor predeter...
3	name	VARCHAR	10000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	NULL
4	value	VARCHAR	10000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	NULL

- **Tabla de historial:**

Esta tabla representa el historial de navegación del usuario se guardan todas las páginas visitadas por el usuario así como las fechas en las cuales accedió a ellas. Esta tabla solo se rellena una sola vez al instalar el plugin en el navegador ya que posteriormente la información se irá recogiendo a través de la tabla de páginas mencionada anteriormente.

#	Nombre	Tipo de datos	Longitud/Conjunto	Sin signo	Permitir NULL	Rellenar con ceros	Predeterminado
1	id	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREMENT
2	iduser	INT	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0
3	url	VARCHAR	10000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	NULL
4	date	DATETIME		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	NULL

- **Tabla de publicidad:**

Las páginas visitadas por el usuario pueden contener enlaces que hacen referencias a mensajes publicitarios. Todos estos enlaces se guardan en esta tabla.

El campo url_actual almacena información referente a la página que nos muestra la publicidad, como podéis ver existe también un campo denominado url_final, esto se debe a que en ocasiones la publicidad no viene dada directamente por el anunciante sino que lo hace a través de terceros, este campo es el que nos dará la URL final después de pasar por todos y cada uno de los intermediarios.

#	Nombre	Tipo de datos	Longitud/Conjunto	Sin signo	Permitir NULL	Rellenar con ceros	Predeterminado
1	id	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREMENT
2	id_page	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	NULL
3	url_actual	VARCHAR	10000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	NULL
4	url_final	VARCHAR	10000	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	NULL

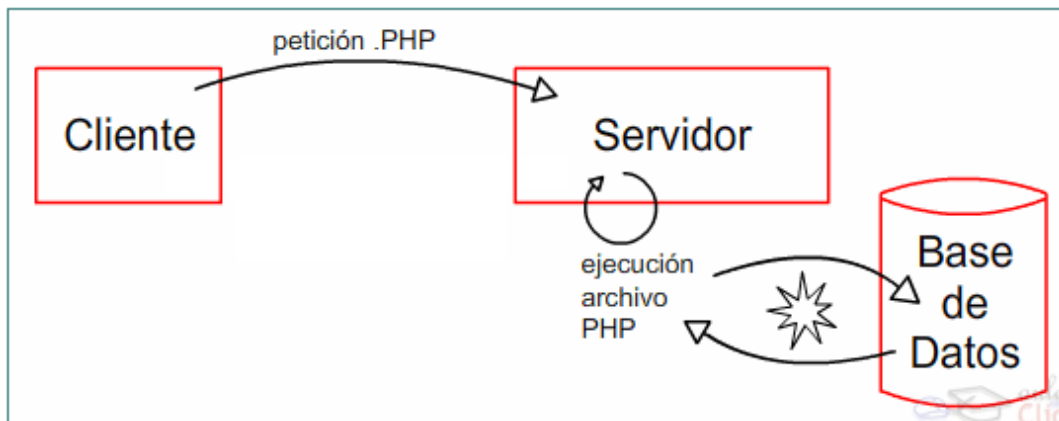
3.1.4 PHP

PHP es un lenguaje de programación de uso general que se ejecuta en el lado del servidor creado principalmente para el desarrollo de contenidos web dinámicos. Fue uno de los primeros lenguaje de programación cuyo código se podía incrustar en el documento HTML, dicho código es interpretado por el servidor que con ayuda del procesador de PHP genera la página web resultante.

PHP significa Hypertext Pre-procesor, fue creado originalmente por Ramus Lerdof. Actualmente la implementación del lenguaje está en manos de The PHP Group.

La ventaja de PHP es que puede ser desplegado en la mayoría de servidores web y en casi todos los sistemas operativos. Su gran parecido con otros lenguajes más comunes de programación estructurada nos permite a la mayoría de programadores crear aplicaciones con una curva de aprendizaje muy corta.

El proyecto utiliza lenguaje PHP para conectarse con la base de datos y realizar las operaciones de inserción en las tablas mencionadas anteriormente. PHP proporciona una interfaz de acceso a la base de datos a través de una serie de funciones de la librería de MySQL. Dichas funciones realizan la conexión y desconexión de la base de datos así como la ejecución de sentencias.



Existen varias maneras de conectarse a la base de datos, pero las tres formas más populares de conectarse desde PHP a una base de datos de MySQL son: MySQL, MySQLi y PDO (PHP Data Objects).

La que se utiliza para el proyecto es la MySQLi que es la nueva versión de MySQL orientada a objetos.

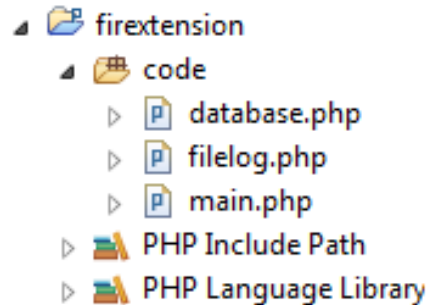
3.1.5 ECLIPSE

Eclipse es un entorno de desarrollo integrado (IDE), de código abierto y multiplataforma. Está escrito principalmente en Java y su uso principal es para el desarrollo de aplicaciones Java, pero también se puede utilizar para desarrollar aplicaciones en otros lenguajes de programación mediante el uso de plugins, incluyendo: Ada, ABAP, C, C ++, COBOL, Fortran, Haskell, JavaScript, Lasso, Lua, Natural, Perl, PHP, Prolog, Python, etc.

En este proyecto fin de carrera hemos utilizado dos versiones, una para el desarrollo de aplicaciones en PHP y otra para el desarrollo de aplicaciones en java.

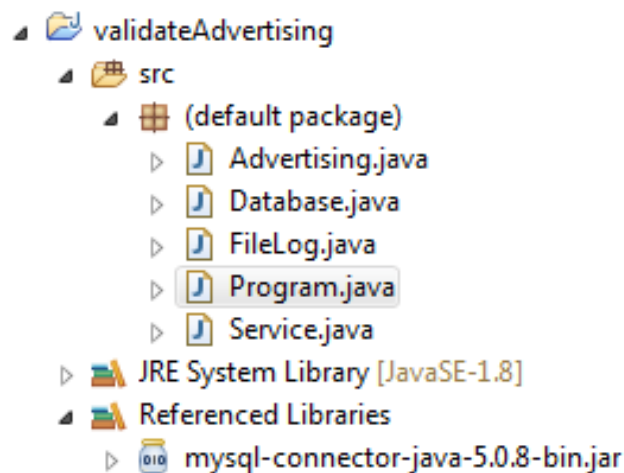


- **Estructuración del código PHP** que realiza las operaciones de inserción de las páginas en la base de datos del servidor.



Las funciones que nos permiten acceder a los datos situados en la base de datos están alojadas en la librería de PHP, la cual se incluye en el proyecto a través de PHP Language Library.

- **Estructuración del código Java** que forma el proceso encargado de recorrer las páginas relativas a la publicidad y obtener su dirección final después de pasar por todas las páginas intermedias.



Las librerías que nos aparecen en el proyecto java de eclipse son las librerías de JRE (Java Runtime Environment), y las librerías proporcionadas por mysql-connector-java-5.0.8-bin.jar la cual nos permitirá acceder a las funciones que trabajan con la base de datos MySQL.

3.1.6 JAVA

Java es un lenguaje de programación orientado a objetos y multiplataforma, al ser multiplataforma permite que el código pueda ser ejecutado en cualquier sistema operativo sin necesidad de que vuelva a ser compilado. Java se ha convertido a partir de 2012 en uno de los lenguajes más populares en uso, especialmente en las plataformas cliente-servidor.

El lenguaje java fue originalmente desarrollado por James Gosling de Sun Microsystem, empresa que fue comprada posteriormente por Oracle, se publico en 1995 como un componente principal de la plataforma Java de Sun Microsystem, la sintaxis del lenguaje deriva de C y C++, sin embargo tiene menos utilidades de bajo nivel que cualquiera de ellas. Las aplicaciones de java son generalmente compiladas en bytecode (clase java) que puede ejecutarse en cualquier maquina virtual java independientemente de la plataforma.

La característica de ser multiplataforma es la que nos ha llevado a seleccionar este lenguaje para realizar el proceso de validación de los diferentes enlaces publicitarios que aparecen en cada una de las páginas recogidas en la base de datos.

Para poder acceder a los datos contenidos en la base de datos de MySQL se ha utilizado un conector (mysql-connector-java-5.0.8-bin.jar), dicho conector nos proporciona una serie de clases con sus correspondientes métodos, los cuales nos permitirán obtener la información referente a las URLs de la tabla advertising.

La clase Database.java es la que se comunica con la base de datos a través del conector, en esta clase se recogen todas los enlaces publicitarios que no hayan sido analizados con anterioridad. Estos enlaces publicitarios recogidos no son más que objetos de la clase Advertising.java.

Cada uno de esos enlaces publicitarios pueden redirigirnos a varias páginas hasta llegar finalmente a la página que nos muestra la publicidad, esto se debe a que muchos de estos sitios se publicitan a través de terceros. El objetivo del proceso es obtener la página final una vez que se haya pasado por todos y cada uno de los intermediarios.

Un ejemplo de intermediarios es googleadservices, el cual es un dominio utilizado por Google Ad Services empresa publicitaria que utiliza la información de navegación del usuario, cookies y otras tecnologías. Con toda esta información googleadservices construye un perfil de usuario el cual puede ser vendido a otros comerciantes y anunciantes.

La clase que nos proporciona los métodos que nos permitirán acceder a la URL final del anunciante se denomina URLConnection.

3.2 Herramientas y tecnologías utilizadas en el cliente

3.2.1 FIREFOX

Mozilla Firefox (llamado simplemente Firefox) es un navegador web libre y de código abierto desarrollado para Microsoft Windows, Android, OS X y GNU/Linux coordinado por la Corporación Mozilla y la Fundación Mozilla. Usa el motor Gecko para renderizar páginas webs, el cual implementa actuales y futuros estándares web.

Firefox comenzó como una rama experimental del proyecto Mozilla a cargo de Dave Hyatt, Joe Hewitt y Blake Ross. A su juicio, las exigencias comerciales del patrocinio de Netscape y el gran número de características de Mozilla Application Suite comprometían la utilidad de este. Para combatir lo que ellos denominaban inflada Mozilla Application Suite, crearon un navegador independiente con la intención de reemplazarla. El 3 de abril de 2003, la Organización Mozilla anuncia que centrarán sus esfuerzos en Firefox y Thunderbird.

El plugin que recogera la información de navegación del usuario está desarrollado para este navegador web.

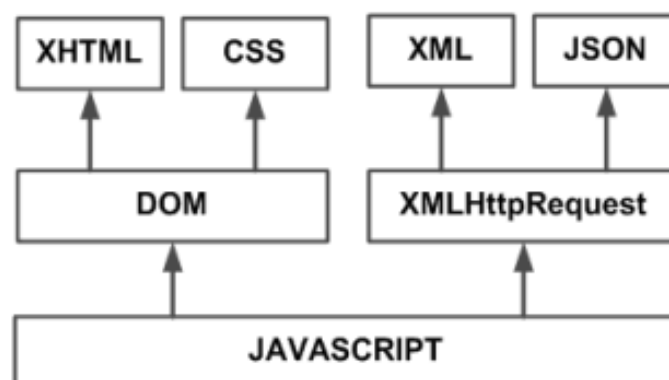
3.2.2 AJAX

El término AJAX se presentó por primera vez en el artículo "[Ajax: A New Approach to Web Applications](#)" publicado por Jesse James Garrett el 18 de Febrero de 2005. Hasta ese momento, no existía un término normalizado que hiciera referencia a un nuevo tipo de aplicación web que estaba apareciendo.

En realidad, el término AJAX es un acrónimo de *Asynchronous JavaScript + XML*, que se puede traducir como "JavaScript asíncrono + XML".

Las tecnologías que forman AJAX son:

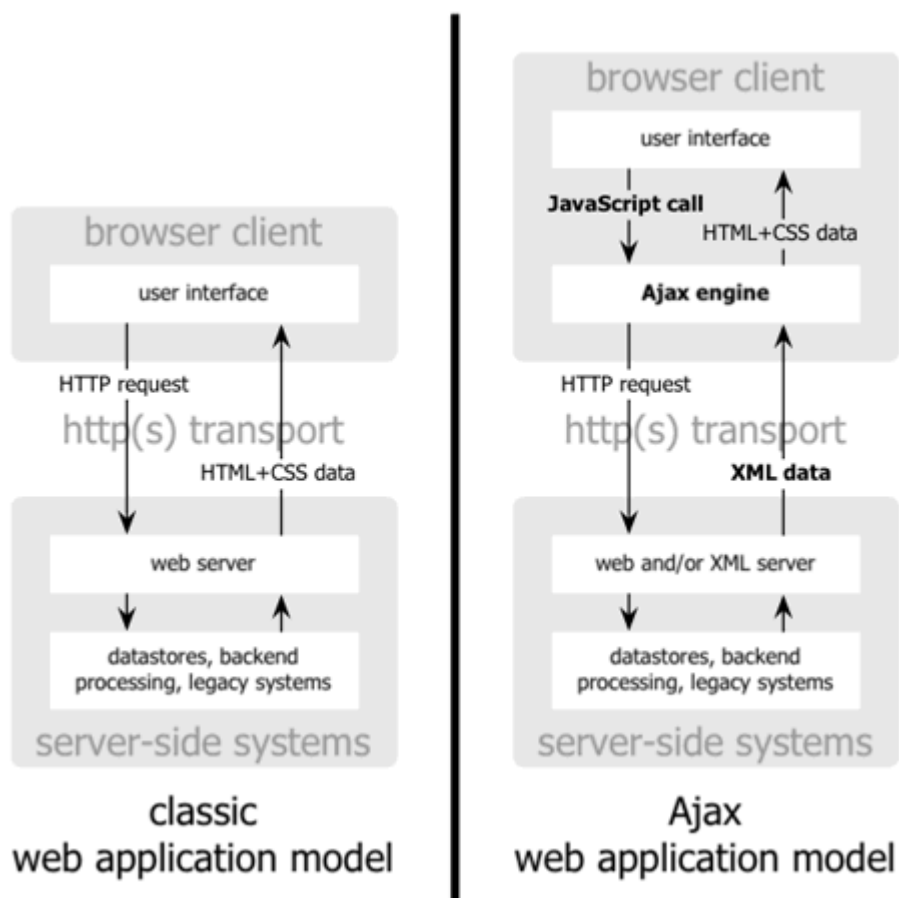
- XHTML y CSS, para crear una presentación basada en estándares.
- DOM, para la interacción y manipulación dinámica de la presentación.
- XML, XSLT y JSON, para el intercambio y la manipulación de información.
- XMLHttpRequest, para el intercambio asíncrono de información.
- JavaScript, para unir todas las demás tecnologías.



Desarrollar aplicaciones AJAX requiere un conocimiento avanzado de todas y cada una de las tecnologías anteriores.

En las aplicaciones web tradicionales, las acciones del usuario en la página (pinchar en un botón, seleccionar un valor de una lista, etc.) desencadenan llamadas al servidor. Una vez procesada la petición del usuario, el servidor devuelve una nueva página HTML al navegador del usuario.

En el siguiente esquema, la imagen de la izquierda muestra el modelo tradicional de las aplicaciones web. La imagen de la derecha muestra el nuevo modelo propuesto por AJAX:



Esta técnica tradicional para crear aplicaciones web funciona correctamente, pero no crea una buena sensación al usuario. Al realizar peticiones continuas al servidor, el usuario debe esperar a que se recargue la página con los cambios solicitados. Si la aplicación debe realizar peticiones continuas, su uso se convierte en algo molesto.

AJAX permite mejorar completamente la interacción del usuario con la aplicación, evitando las recargas constantes de la página, ya que el intercambio de información con el servidor se produce en un segundo plano.

Las aplicaciones construidas con AJAX eliminan la recarga constante de páginas mediante la creación de un elemento intermedio entre el usuario y el servidor. La nueva capa intermedia de AJAX mejora la respuesta de la aplicación, ya que el usuario nunca se encuentra con una ventana del navegador vacía esperando la respuesta del servidor.

El proyecto utiliza las llamadas AJAX para enviar al servidor de PHP la información referente a la navegación del usuario, estas llamadas se realizan de manera asíncrona de tal forma que el proceso sea transparente para el usuario que está utilizando el navegador.

3.2.3 JAVASCRIPT

Javascript es un lenguaje de programación interpretado, se define orientado a objetos, basado en prototipos y débilmente tipado. Este lenguaje se utiliza principalmente del lado del cliente implementado como parte del navegador web permitiendo mejoras en el interfaz de usuario de la aplicación y páginas web dinámicas. Todos los navegadores web interpretan el javascript integrado en las páginas.

La interacción de javascript con la página web se realiza a través del DOM (Document Object Model), mientras que el intercambio de información con el servidor PHP se emplea el formato Json.

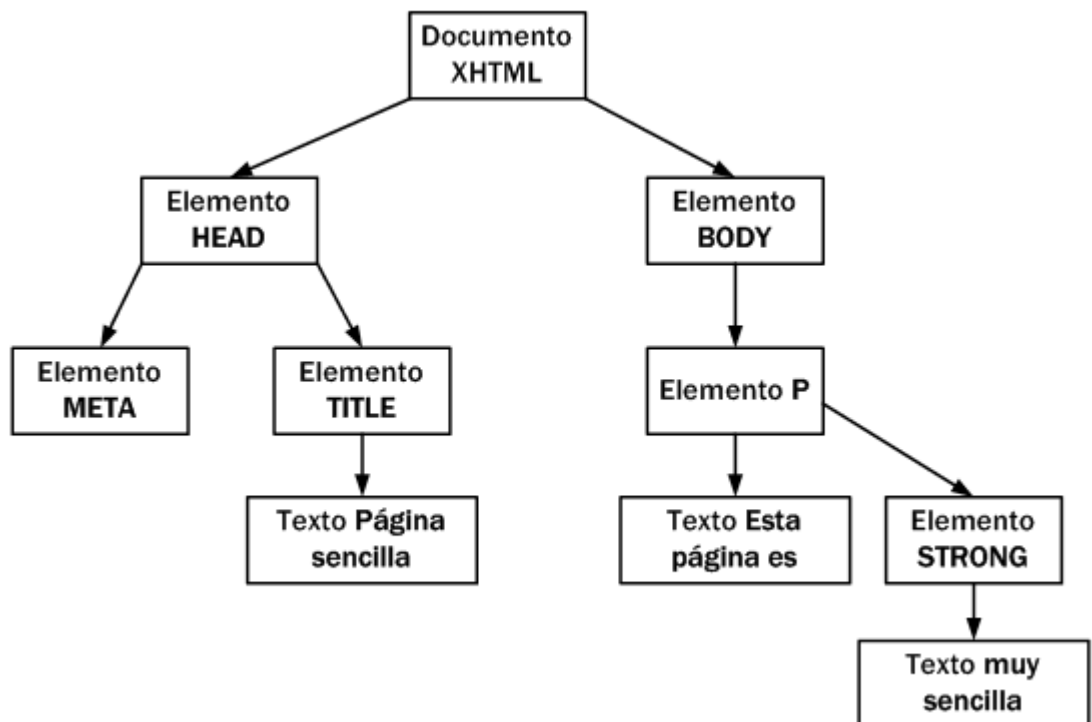
- **DOM** o *Document Object Model* es un conjunto de utilidades específicamente diseñadas para manipular documentos XML. Por extensión, DOM también se puede utilizar para manipular documentos XHTML y HTML. Técnicamente, DOM es una API de funciones que se pueden utilizar para manipular las páginas XHTML de forma rápida y eficiente.

Antes de poder utilizar sus funciones, DOM transforma internamente el archivo XML original en una estructura más fácil de manejar formada por una jerarquía de nodos. De esta forma, DOM transforma el código XML en una serie de nodos interconectados en forma de *árbol*.

El árbol generado no sólo representa los contenidos del archivo original (mediante los nodos del árbol) sino que también representa

sus relaciones (mediante las ramas del árbol que conectan los nodos).

Antes de poder utilizar las funciones de DOM, los navegadores convierten automáticamente la página HTML anterior en la siguiente estructura de árbol de nodos:



El proyecto utiliza el lenguaje javascript, con el objetivo de poder acceder a los diferentes elementos de la página a través del API de DOM. Por cada uno de los elementos de la página se analizan sus atributos los cuales contienen información a cerca de la publicidad incluida en la página visitada.

- **JSON**, acrónimo de *JavaScript Object Notation*, es un formato ligero para el intercambio de datos. JSON es un subconjunto de la notación literal de objetos de JavaScript que no requiere el uso de XML.

La simplicidad de JSON ha dado lugar a la generalización de su uso, especialmente como alternativa a XML en AJAX. Una de las

supuestas ventajas de JSON sobre XML como formato de intercambio de datos en este contexto es que es mucho más sencillo escribir un analizador sintáctico (parser) de JSON. En JavaScript, un texto JSON se puede analizar fácilmente usando la función `eval()`, lo cual ha sido fundamental para que JSON haya sido aceptado por parte de la comunidad de desarrolladores AJAX, debido a la ubicuidad de JavaScript en casi cualquier navegador web.

En el proyecto se ha utilizado el formato JSON para facilitar el intercambio de información entre el código javascript del navegador del cliente y el código PHP implementado en el servidor, este intercambio se realiza a través de AJAX. Esto es posible gracias a las librerías que nos proporciona PHP (`json_decode`) y que nos permite extraer la información en dicho formato y asignarla a una variable PHP.

En cada llamada AJAX enviamos la información de la página visitada, es decir, el usuario que ha visitado la página, las cookies de dicha página y los diferentes enlaces publicitarios que aparecen. Toda esta jerarquía de objetos javascript se transforma al formato JSON que posteriormente será interpretado por PHP.

3.2.4 EXTENSIONES

Las extensiones son pequeños programas que se añaden al navegador para adaptarlo a tus necesidades y de esta forma personalizar tu experiencia de navegación. Gracias a las extensiones podemos incluir nuevas funcionalidades al navegador, y a todos sus componentes. Podemos interactuar con las pestañas abiertas del navegador así como analizar el contenido de las páginas.

La comunidad de desarrollo de mozilla nos ofrece tres opciones para desarrollar extensiones:

- **Add-on sdk extensión:** Estas extensiones se crean utilizando una serie de librerías de alto nivel de javascript y además no requiere reiniciar el navegador cada vez que instalemos la extensión desarrollada. Las librerías nos permitirán interactuar con el navegador y con el contenido de sus páginas. Con add-on sdk tu puedes crear extensiones utilizando solamente javascript, html y css.

- **Restarless:** Las librerías que se utilizan para crear este tipo de extensiones son de bajo nivel, al igual que en el caso anterior las extensión restarless no necesitan reiniciar el navegador para cada vez que se instalan.
- **Legacy extensions:** Antes de Firefox 4 y del motor de gecko 2 que lo hace funcionar, esta era la única manera de desarrollar extensiones. Esta metodología ha sido ampliamente superada por las 2 anteriores, las cuales se desarrollan en una capa superior, sin embargo todas las librerías de legacy se pueden utilizar en las capas superiores (restarless y add-on).

En el desarrollo de proyecto se ha optado por utilizar Add-on SDK, ya que las librerías de alto nivel que posee resuelven las necesidades de la aplicación, abstrayéndose de realizar ciertos controles de los cuales se encargan las capas inferiores. Otros de los factores que han llevado a su elección ha sido su fácil instalación.

Para recoger la información de navegación del usuario se ha empleado la librería de alto nivel denominada Tabs, para la gestión del historial se ha utilizado Places y el almacenamiento del identificador del usuario se realiza a través de la librería de bajo nivel denominada io/text-stream, todas estas librerías se pueden ver con más detalle en el anexo 4 “Librerías utilizadas en el desarrollo del Add-on.”

4. RESULTADOS

Una vez que se ha implementado el plugin deberemos comprobar su correcto funcionamiento a través de los casos de uso y las baterías de pruebas que determinarán si se está almacenando de manera correcta toda la información necesaria para el estudio.

4.1 Caso de uso










En este apartado se detalla el caso de uso del aplicativo. Un caso de uso es una descripción de los pasos o las actividades que deberán realizarse para llevar a cabo algún proceso.

En el caso de este proyecto solo tenemos un caso de uso de la aplicación, el cual consiste en abrir una ventana o pestaña del navegador Firefox y navegar por las distintas páginas webs según los gustos e inquietudes de los usuarios.

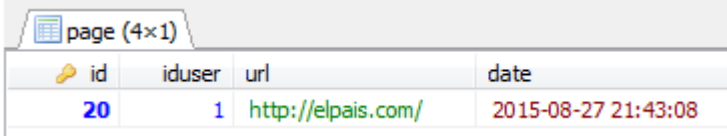
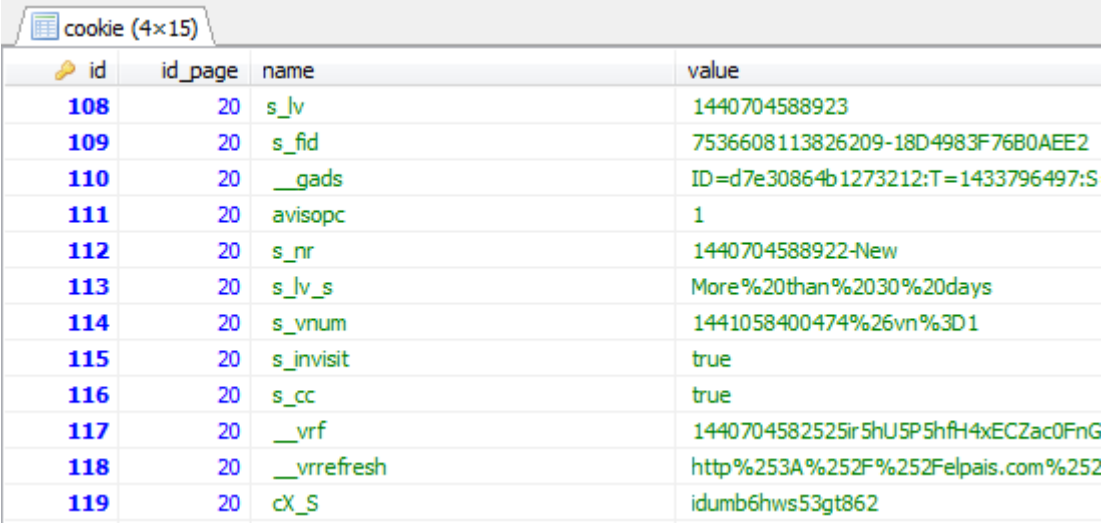
4.2 Batería de pruebas

En este apartado se mostrarán las baterías de pruebas realizadas sobre el aplicativo. Cada caso de test está identificado por su id (TC-n) y su nombre.

TC.1. Carga del historial de navegación

ELEMENTO	DETALLES																						
Nombre	Carga del historial de navegación del usuario al instalar el plugin																						
Participantes	Usuario a través del navegador																						
Descripción	El usuario abre las opciones del navegador e instala el correspondiente plugin																						
Precondiciones	El actor utiliza el navegador Firefox y tiene conectividad con la red																						
Resultado	Una vez instalado el plugin el cliente envía la información del usuario y de su historial a la base de datos del servidor y lo almacena en las tablas user e history.																						
Evidencias	<p>Tabla user:</p> <table><tr><th> id</th><th> code_hex</th><th>date</th></tr><tr><td>1</td><td>58eb1d52-f2a2-4d70-9007-e3bc3d15222e</td><td>2015-08-27 20:39:47</td></tr></table> <p>Tabla history:</p> <table><tr><th> id</th><th>iduser</th><th>url</th><th>date</th></tr><tr><td>1</td><td>1</td><td>https://www.mozilla.org/en-US/firefox/38.0.1/firstrun/</td><td>2015-06-01 23:08:25</td></tr><tr><td>2</td><td>1</td><td>https://www.google.es/search?q=mdn&ie=utf-8&oe=utf...</td><td>2015-06-01 23:08:32</td></tr><tr><td>3</td><td>1</td><td>https://developer.mozilla.org/es/</td><td>2015-06-01 23:08:38</td></tr></table>	 id	 code_hex	date	1	58eb1d52-f2a2-4d70-9007-e3bc3d15222e	2015-08-27 20:39:47	 id	iduser	url	date	1	1	https://www.mozilla.org/en-US/firefox/38.0.1/firstrun/	2015-06-01 23:08:25	2	1	https://www.google.es/search?q=mdn&ie=utf-8&oe=utf...	2015-06-01 23:08:32	3	1	https://developer.mozilla.org/es/	2015-06-01 23:08:38
 id	 code_hex	date																					
1	58eb1d52-f2a2-4d70-9007-e3bc3d15222e	2015-08-27 20:39:47																					
 id	iduser	url	date																				
1	1	https://www.mozilla.org/en-US/firefox/38.0.1/firstrun/	2015-06-01 23:08:25																				
2	1	https://www.google.es/search?q=mdn&ie=utf-8&oe=utf...	2015-06-01 23:08:32																				
3	1	https://developer.mozilla.org/es/	2015-06-01 23:08:38																				

TC.2. Almacenamiento de la URL y sus cookies

ELEMENTO	DETALLES
Nombre	Testeo del almacenamiento de la URL de la página y sus correspondientes cookies
Participantes	Usuario a través del navegador
Descripción	El usuario introduce una URL en la barra de direcciones del navegador
Precondiciones	El actor utiliza el navegador Firefox y tiene conectividad con la red
Resultado	Una vez cargada la página en el navegador el plugin envía información referente a la URL visitada y las cookies encontradas en dicha página
Evidencias	<p>Tabla de páginas:</p>  <p>Tabla cookies:</p> 

TC.3. Almacenamiento de los enlaces publicitarios

ELEMENTO	DETALLES																																																																																																				
Nombre	Testeo del almacenamiento de la URL de la página y correspondientes enlaces publicitarios																																																																																																				
Participantes	Usuario a través del navegador																																																																																																				
Descripción	El usuario introduce una URL en la barra de direcciones del navegador																																																																																																				
Precondiciones	El actor utiliza el navegador Firefox y tiene conectividad con la red																																																																																																				
Resultado	Una vez cargada la página en el navegador el plugin envía información referente a la URL visitada y los enlaces publicitarios encontrados en dicha página																																																																																																				
Evidencias	<p>Tabla advertising:</p> <table><tr><th>▲ id</th><th>▲ id_page</th><th>url_actual</th><th>url_final</th></tr><tr><td>1</td><td>33</td><td>http://cookies.unidadeditorial.es/</td><td>(NULL)</td></tr><tr><td>2</td><td>33</td><td>http://www.tiramillas.net/?cid=MENUPROD35603&s_kw...</td><td>(NULL)</td></tr><tr><td>3</td><td>33</td><td>http://www.tecnopasion.com/?cid=MENUPROD35603&s_...</td><td>(NULL)</td></tr><tr><td>4</td><td>33</td><td>http://es.ign.com/?cid=MENUPROD35603&s_kw=ign</td><td>(NULL)</td></tr><tr><td>5</td><td>33</td><td>http://es.ign.com/</td><td>(NULL)</td></tr><tr><td>6</td><td>33</td><td>http://es.ign.com/</td><td>(NULL)</td></tr><tr><td>7</td><td>33</td><td>http://es.ign.com/</td><td>(NULL)</td></tr><tr><td>8</td><td>33</td><td>http://es.ign.com/warcraft-1/94942/news/el-trailer-de-w...</td><td>(NULL)</td></tr><tr><td>9</td><td>33</td><td>http://es.ign.com/everest/96043/news/everest-spots-pr...</td><td>(NULL)</td></tr><tr><td>10</td><td>33</td><td>http://es.ign.com/everest/96043/news/everest-spots-pr...</td><td>(NULL)</td></tr><tr><td>11</td><td>33</td><td>http://es.ign.com/everest/96043/news/everest-spots-pr...</td><td>(NULL)</td></tr><tr><td>12</td><td>33</td><td>http://es.ign.com/final-fantasy-versus-xiii-ps4/96026/ne...</td><td>(NULL)</td></tr><tr><td>13</td><td>33</td><td>http://es.ign.com/cine/96023/news/muere-wes-craven-a...</td><td>(NULL)</td></tr><tr><td>14</td><td>33</td><td>http://es.ign.com/daily-fix/95379/fix/comics-cine-juegos-...</td><td>(NULL)</td></tr><tr><td>15</td><td>33</td><td>http://www.tiramillas.net/2015/09/01/tendencias/14410...</td><td>(NULL)</td></tr><tr><td>16</td><td>33</td><td>http://www.tiramillas.net/2015/09/01/tendencias/14410...</td><td>(NULL)</td></tr><tr><td>17</td><td>33</td><td>http://www.tecnopasion.com/</td><td>(NULL)</td></tr><tr><td>18</td><td>33</td><td>http://www.tecnopasion.com/crean-la-primera-bicideta-...</td><td>(NULL)</td></tr><tr><td>19</td><td>33</td><td>http://www.tecnopasion.com/crean-la-primera-bicideta-...</td><td>(NULL)</td></tr><tr><td>20</td><td>33</td><td>http://www.tecnopasion.com/</td><td>(NULL)</td></tr><tr><td>21</td><td>33</td><td>http://www.tecnopasion.com/el-samsung-gear-s2-podria...</td><td>(NULL)</td></tr><tr><td>22</td><td>33</td><td>http://www.tecnopasion.com/el-samsung-gear-s2-podria...</td><td>(NULL)</td></tr><tr><td>23</td><td>33</td><td>http://www.elmundo.es/?cid=SIN12201</td><td>(NULL)</td></tr><tr><td>24</td><td>33</td><td>http://www.elmundo.es/enredados/2015/08/31/55e47f6...</td><td>(NULL)</td></tr></table>	▲ id	▲ id_page	url_actual	url_final	1	33	http://cookies.unidadeditorial.es/	(NULL)	2	33	http://www.tiramillas.net/?cid=MENUPROD35603&s_kw...	(NULL)	3	33	http://www.tecnopasion.com/?cid=MENUPROD35603&s_...	(NULL)	4	33	http://es.ign.com/?cid=MENUPROD35603&s_kw=ign	(NULL)	5	33	http://es.ign.com/	(NULL)	6	33	http://es.ign.com/	(NULL)	7	33	http://es.ign.com/	(NULL)	8	33	http://es.ign.com/warcraft-1/94942/news/el-trailer-de-w...	(NULL)	9	33	http://es.ign.com/everest/96043/news/everest-spots-pr...	(NULL)	10	33	http://es.ign.com/everest/96043/news/everest-spots-pr...	(NULL)	11	33	http://es.ign.com/everest/96043/news/everest-spots-pr...	(NULL)	12	33	http://es.ign.com/final-fantasy-versus-xiii-ps4/96026/ne...	(NULL)	13	33	http://es.ign.com/cine/96023/news/muere-wes-craven-a...	(NULL)	14	33	http://es.ign.com/daily-fix/95379/fix/comics-cine-juegos-...	(NULL)	15	33	http://www.tiramillas.net/2015/09/01/tendencias/14410...	(NULL)	16	33	http://www.tiramillas.net/2015/09/01/tendencias/14410...	(NULL)	17	33	http://www.tecnopasion.com/	(NULL)	18	33	http://www.tecnopasion.com/crean-la-primera-bicideta-...	(NULL)	19	33	http://www.tecnopasion.com/crean-la-primera-bicideta-...	(NULL)	20	33	http://www.tecnopasion.com/	(NULL)	21	33	http://www.tecnopasion.com/el-samsung-gear-s2-podria...	(NULL)	22	33	http://www.tecnopasion.com/el-samsung-gear-s2-podria...	(NULL)	23	33	http://www.elmundo.es/?cid=SIN12201	(NULL)	24	33	http://www.elmundo.es/enredados/2015/08/31/55e47f6...	(NULL)
▲ id	▲ id_page	url_actual	url_final																																																																																																		
1	33	http://cookies.unidadeditorial.es/	(NULL)																																																																																																		
2	33	http://www.tiramillas.net/?cid=MENUPROD35603&s_kw...	(NULL)																																																																																																		
3	33	http://www.tecnopasion.com/?cid=MENUPROD35603&s_...	(NULL)																																																																																																		
4	33	http://es.ign.com/?cid=MENUPROD35603&s_kw=ign	(NULL)																																																																																																		
5	33	http://es.ign.com/	(NULL)																																																																																																		
6	33	http://es.ign.com/	(NULL)																																																																																																		
7	33	http://es.ign.com/	(NULL)																																																																																																		
8	33	http://es.ign.com/warcraft-1/94942/news/el-trailer-de-w...	(NULL)																																																																																																		
9	33	http://es.ign.com/everest/96043/news/everest-spots-pr...	(NULL)																																																																																																		
10	33	http://es.ign.com/everest/96043/news/everest-spots-pr...	(NULL)																																																																																																		
11	33	http://es.ign.com/everest/96043/news/everest-spots-pr...	(NULL)																																																																																																		
12	33	http://es.ign.com/final-fantasy-versus-xiii-ps4/96026/ne...	(NULL)																																																																																																		
13	33	http://es.ign.com/cine/96023/news/muere-wes-craven-a...	(NULL)																																																																																																		
14	33	http://es.ign.com/daily-fix/95379/fix/comics-cine-juegos-...	(NULL)																																																																																																		
15	33	http://www.tiramillas.net/2015/09/01/tendencias/14410...	(NULL)																																																																																																		
16	33	http://www.tiramillas.net/2015/09/01/tendencias/14410...	(NULL)																																																																																																		
17	33	http://www.tecnopasion.com/	(NULL)																																																																																																		
18	33	http://www.tecnopasion.com/crean-la-primera-bicideta-...	(NULL)																																																																																																		
19	33	http://www.tecnopasion.com/crean-la-primera-bicideta-...	(NULL)																																																																																																		
20	33	http://www.tecnopasion.com/	(NULL)																																																																																																		
21	33	http://www.tecnopasion.com/el-samsung-gear-s2-podria...	(NULL)																																																																																																		
22	33	http://www.tecnopasion.com/el-samsung-gear-s2-podria...	(NULL)																																																																																																		
23	33	http://www.elmundo.es/?cid=SIN12201	(NULL)																																																																																																		
24	33	http://www.elmundo.es/enredados/2015/08/31/55e47f6...	(NULL)																																																																																																		

TC.4. Obtención de la URL final tras pasar por los intermediarios

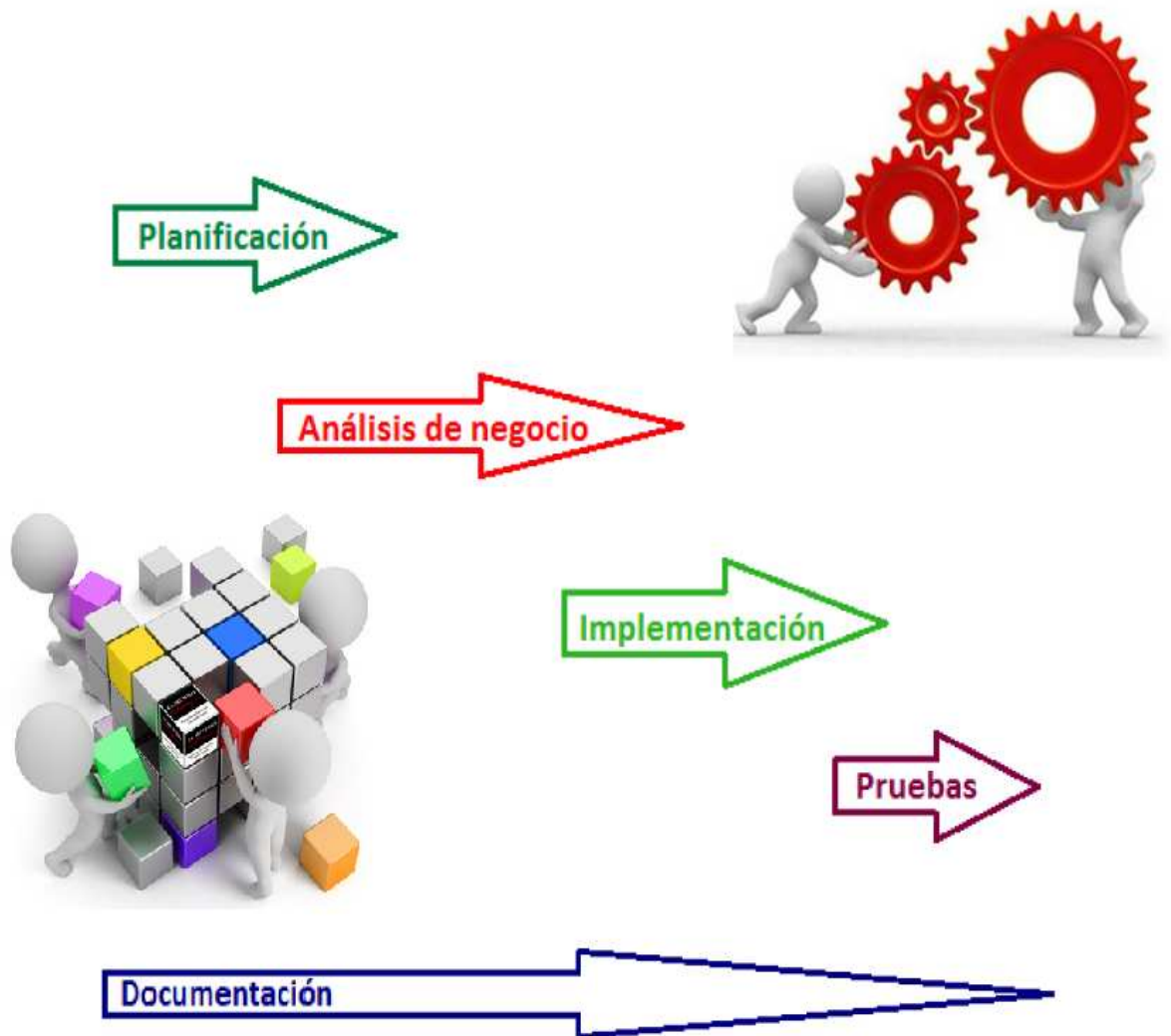
ELEMENTO	DETALLES																																																																																																				
Nombre	Testeo de la obtención de la URL final después de pasar por los diferentes intermediarios																																																																																																				
Participantes	El servicio arrancado en el servidor																																																																																																				
Descripción	El servicio accede a todos aquellos enlaces que no hayan sido analizados previamente y obtiene la URL final en caso de que tengan varios intermediarios. Si no tienen intermediarios la URL actual será la final.																																																																																																				
Precondiciones	Tener arrancado el motor de MySQL que nos permitirá realizar las consultas para obtener las URLs que serán procesadas y tener las tablas creadas previamente.																																																																																																				
Resultado	El servicio recorrerá la tabla de advertising y por cada enlace publicitario irá rellenado el campo de URL_final.																																																																																																				
Evidencias	<p>Tabla advertising:</p> <table><tr><th>▲ id</th><th>▲ id_page</th><th>url_actual</th><th>url_final</th></tr><tr><td>1</td><td>33</td><td>http://cookies.unidadeditorial.es/</td><td>http://cookies.unidadeditorial.es/</td></tr><tr><td>2</td><td>33</td><td>http://www.tiramillas.net/?cid=MEN...</td><td>http://www.tiramillas.net/?cid=MEN...</td></tr><tr><td>3</td><td>33</td><td>http://www.tecnopasion.com/?cid=...</td><td>http://www.tecnopasion.com/?cid=...</td></tr><tr><td>4</td><td>33</td><td>http://es.ign.com/?cid=MENUPROD...</td><td>http://es.ign.com/?cid=MENUPROD...</td></tr><tr><td>5</td><td>33</td><td>http://es.ign.com/</td><td>http://es.ign.com/</td></tr><tr><td>6</td><td>33</td><td>http://es.ign.com/</td><td>http://es.ign.com/</td></tr><tr><td>7</td><td>33</td><td>http://es.ign.com/</td><td>http://es.ign.com/</td></tr><tr><td>8</td><td>33</td><td>http://es.ign.com/warcraft-1/94942...</td><td>http://es.ign.com/warcraft-1/94942...</td></tr><tr><td>9</td><td>33</td><td>http://es.ign.com/everest/96043/ne...</td><td>http://es.ign.com/everest/96043/n...</td></tr><tr><td>10</td><td>33</td><td>http://es.ign.com/everest/96043/ne...</td><td>http://es.ign.com/everest/96043/n...</td></tr><tr><td>11</td><td>33</td><td>http://es.ign.com/everest/96043/ne...</td><td>http://es.ign.com/everest/96043/n...</td></tr><tr><td>12</td><td>33</td><td>http://es.ign.com/final-fantasy-vers...</td><td>http://es.ign.com/final-fantasy-vers...</td></tr><tr><td>13</td><td>33</td><td>http://es.ign.com/cine/96023/news/...</td><td>http://es.ign.com/cine/96023/news/...</td></tr><tr><td>14</td><td>33</td><td>http://es.ign.com/daily-fix/95379/fi...</td><td>http://es.ign.com/daily-fix/95379/fi...</td></tr><tr><td>15</td><td>33</td><td>http://www.tiramillas.net/2015/09/0...</td><td>http://www.tiramillas.net/2015/09/...</td></tr><tr><td>16</td><td>33</td><td>http://www.tiramillas.net/2015/09/0...</td><td>http://www.tiramillas.net/2015/09/...</td></tr><tr><td>17</td><td>33</td><td>http://www.tecnopasion.com/</td><td>http://www.tecnopasion.com/</td></tr><tr><td>18</td><td>33</td><td>http://www.tecnopasion.com/crean...</td><td>http://www.tecnopasion.com/crean...</td></tr><tr><td>19</td><td>33</td><td>http://www.tecnopasion.com/crean...</td><td>http://www.tecnopasion.com/crean...</td></tr><tr><td>20</td><td>33</td><td>http://www.tecnopasion.com/</td><td>http://www.tecnopasion.com/</td></tr><tr><td>21</td><td>33</td><td>http://www.tecnopasion.com/el-sa...</td><td>http://www.tecnopasion.com/el-sa...</td></tr><tr><td>22</td><td>33</td><td>http://www.tecnopasion.com/el-sa...</td><td>http://www.tecnopasion.com/el-sa...</td></tr><tr><td>23</td><td>33</td><td>http://www.elmundo.es/?cid=SIN12...</td><td>http://www.elmundo.es/?cid=SIN12...</td></tr><tr><td>24</td><td>33</td><td>http://www.elmundo.es/enredados/...</td><td>http://www.elmundo.es/enredados/...</td></tr></table>	▲ id	▲ id_page	url_actual	url_final	1	33	http://cookies.unidadeditorial.es/	http://cookies.unidadeditorial.es/	2	33	http://www.tiramillas.net/?cid=MEN...	http://www.tiramillas.net/?cid=MEN...	3	33	http://www.tecnopasion.com/?cid=...	http://www.tecnopasion.com/?cid=...	4	33	http://es.ign.com/?cid=MENUPROD...	http://es.ign.com/?cid=MENUPROD...	5	33	http://es.ign.com/	http://es.ign.com/	6	33	http://es.ign.com/	http://es.ign.com/	7	33	http://es.ign.com/	http://es.ign.com/	8	33	http://es.ign.com/warcraft-1/94942...	http://es.ign.com/warcraft-1/94942...	9	33	http://es.ign.com/everest/96043/ne...	http://es.ign.com/everest/96043/n...	10	33	http://es.ign.com/everest/96043/ne...	http://es.ign.com/everest/96043/n...	11	33	http://es.ign.com/everest/96043/ne...	http://es.ign.com/everest/96043/n...	12	33	http://es.ign.com/final-fantasy-vers...	http://es.ign.com/final-fantasy-vers...	13	33	http://es.ign.com/cine/96023/news/...	http://es.ign.com/cine/96023/news/...	14	33	http://es.ign.com/daily-fix/95379/fi...	http://es.ign.com/daily-fix/95379/fi...	15	33	http://www.tiramillas.net/2015/09/0...	http://www.tiramillas.net/2015/09/...	16	33	http://www.tiramillas.net/2015/09/0...	http://www.tiramillas.net/2015/09/...	17	33	http://www.tecnopasion.com/	http://www.tecnopasion.com/	18	33	http://www.tecnopasion.com/crean...	http://www.tecnopasion.com/crean...	19	33	http://www.tecnopasion.com/crean...	http://www.tecnopasion.com/crean...	20	33	http://www.tecnopasion.com/	http://www.tecnopasion.com/	21	33	http://www.tecnopasion.com/el-sa...	http://www.tecnopasion.com/el-sa...	22	33	http://www.tecnopasion.com/el-sa...	http://www.tecnopasion.com/el-sa...	23	33	http://www.elmundo.es/?cid=SIN12...	http://www.elmundo.es/?cid=SIN12...	24	33	http://www.elmundo.es/enredados/...	http://www.elmundo.es/enredados/...
▲ id	▲ id_page	url_actual	url_final																																																																																																		
1	33	http://cookies.unidadeditorial.es/	http://cookies.unidadeditorial.es/																																																																																																		
2	33	http://www.tiramillas.net/?cid=MEN...	http://www.tiramillas.net/?cid=MEN...																																																																																																		
3	33	http://www.tecnopasion.com/?cid=...	http://www.tecnopasion.com/?cid=...																																																																																																		
4	33	http://es.ign.com/?cid=MENUPROD...	http://es.ign.com/?cid=MENUPROD...																																																																																																		
5	33	http://es.ign.com/	http://es.ign.com/																																																																																																		
6	33	http://es.ign.com/	http://es.ign.com/																																																																																																		
7	33	http://es.ign.com/	http://es.ign.com/																																																																																																		
8	33	http://es.ign.com/warcraft-1/94942...	http://es.ign.com/warcraft-1/94942...																																																																																																		
9	33	http://es.ign.com/everest/96043/ne...	http://es.ign.com/everest/96043/n...																																																																																																		
10	33	http://es.ign.com/everest/96043/ne...	http://es.ign.com/everest/96043/n...																																																																																																		
11	33	http://es.ign.com/everest/96043/ne...	http://es.ign.com/everest/96043/n...																																																																																																		
12	33	http://es.ign.com/final-fantasy-vers...	http://es.ign.com/final-fantasy-vers...																																																																																																		
13	33	http://es.ign.com/cine/96023/news/...	http://es.ign.com/cine/96023/news/...																																																																																																		
14	33	http://es.ign.com/daily-fix/95379/fi...	http://es.ign.com/daily-fix/95379/fi...																																																																																																		
15	33	http://www.tiramillas.net/2015/09/0...	http://www.tiramillas.net/2015/09/...																																																																																																		
16	33	http://www.tiramillas.net/2015/09/0...	http://www.tiramillas.net/2015/09/...																																																																																																		
17	33	http://www.tecnopasion.com/	http://www.tecnopasion.com/																																																																																																		
18	33	http://www.tecnopasion.com/crean...	http://www.tecnopasion.com/crean...																																																																																																		
19	33	http://www.tecnopasion.com/crean...	http://www.tecnopasion.com/crean...																																																																																																		
20	33	http://www.tecnopasion.com/	http://www.tecnopasion.com/																																																																																																		
21	33	http://www.tecnopasion.com/el-sa...	http://www.tecnopasion.com/el-sa...																																																																																																		
22	33	http://www.tecnopasion.com/el-sa...	http://www.tecnopasion.com/el-sa...																																																																																																		
23	33	http://www.elmundo.es/?cid=SIN12...	http://www.elmundo.es/?cid=SIN12...																																																																																																		
24	33	http://www.elmundo.es/enredados/...	http://www.elmundo.es/enredados/...																																																																																																		

5. PLANIFICACIÓN Y PRESUPUESTO

En este capítulo se detallan las fases de desarrollo de la aplicación, el personal necesario para su realización y la estimación del presupuesto.

5.1 Fases del proyecto

Se ha dividido el proyecto en las siguientes fases.



5.2 Personal

En este apartado se detallan los roles necesarios para el desarrollo del proyecto.

- **Jefe de proyecto:** Encargado de coordinar todas las actividades a realizar y las personas involucradas. Responsable de cumplir con los plazos y el presupuesto.
- **Analista funcional:** No tiene por qué ser alguien experto en informática, ya que su labor principal es conocer el negocio y tomar requisitos teniendo en cuenta las necesidades del usuario.
- **Programador:** Una vez terminada la fase de diseño, el programador se encarga de darle forma a todos los requisitos de negocio y técnicos tomados previamente.
- **Probador de software:** Su cometido es planificar una batería de pruebas del software, detectar errores y comunicarlos para su resolución. También sugieren mejoras a raíz de la evaluación realizada.

Cabe reseñar que la mayoría de las veces que se desarrolla un proyecto, una persona realiza más de un rol. No es raro encontrarse el papel de analista con los roles de programador y probador de software

Fase	Puesto
Planificación	Jefe de proyecto
Análisis de negocio	Analista funcional
Implementación	Programador
Pruebas	Probador de software
Documentación	Jefe de proyecto Analista funcional Programador Probador de software

5.3 Presupuesto

En este apartado se realizará una estimación del presupuesto para la realización del proyecto.

5.3.1 ESTIMACIÓN DE TIEMPOS

La planificación está realizada en base a una jornada de 8 horas.

Fase	Jornadas
Planificación	6
Análisis de negocio	7
Implementación	20
Pruebas	7
Documentación	20

Tiempo estimado total: 60 jornadas / 480 horas

5.3.2 ESTIMACIÓN DE COSTES

Fase	Puesto	Coste hora (EUR)	Horas	Coste total (EUR)
Planificación	Jefe de Proyecto	50	48	2400
Análisis de negocio	Analista funcional	35	56	1960
Implementación	Programador	25	160	4000
Pruebas	Probador de software	20	56	1120
Documentación	Jefe de proyecto Analista funcional Programador Probador de software	15	160	2400

Coste estimado total: 11.880 euros

6. CONCLUSIONES Y FUTUROS DESARROLLOS

6.1 Conclusiones

La principal conclusión que se puede extraer es que se ha cumplido el objetivo primordial del proyecto, que consistía en recoger la mayor cantidad de información posible acerca de la navegación de los usuarios a través de internet.

El objetivo se encuentra dividido en dos partes bien diferenciadas:

1. El análisis de todas y cada una de las páginas html visitadas en busca de enlaces publicitarios.
2. Almacenamiento de la información de cada uno de los usuarios en una base de datos para su posterior estudio.

- **Conclusiones personales**

Como principal conclusión, a nivel personal ha sido un proyecto enriquecedor desde el punto de vista académico.

La fase de investigación me ha permitido conocer las diferentes herramientas que existen actualmente relacionadas con el tema de la publicidad, y como a través de ellas podemos evitar el seguimiento que realizan las empresas publicitarias de nuestra navegación, además de evitar el seguimiento nos ofrece la posibilidad de consultar sus políticas de privacidad y así saber si dicha empresa está vulnerando nuestra intimidad.

6.2 Futuros desarrollos

6.2.1 CLASIFICACIÓN DE LAS PÁGINAS VISITADAS

Actualmente en la base de datos están almacenadas todas las páginas visitadas por los usuarios, como se puede observar aparecen páginas de muy diversos temas (política, deportes, cultura, alimentación, ocio, salud), la idea es recorrer las páginas visitadas y realizar una clasificación según el tema, de esta forma se podrá conocer cuáles son los gustos e inquietudes del usuario.

6.2.2 INTEGRACIÓN DEL PLUGIN EN UNA PLATAFORMA CROWDSOURCING

Integración del plugin en una plataforma abierta a un número indeterminado de personas para el análisis de datos a gran escala. Es decir, una vez que tengamos el plugin instalado en un gran número de usuarios empezaremos a recibir información de todos ellos de manera que podamos hacer un análisis de los datos para inferir patrones de personalización en la publicidad mucho mejores que lo que podamos realizar con datos de un único usuario.

6.3 Dificultades encontradas

A lo largo de la realización se han encontrado distintas dificultades. A continuación se pasan a explicar las más reseñables.

- **Anuncios a través de terceros**

Muchos de los enlaces publicitarios que encontramos en las páginas no hacen referencia directa a la empresa que nos muestra la publicidad, sino que utilizan una serie de intermediarios, esto era un problema ya que la URL almacenada era la del intermediario y no la del anunciante.



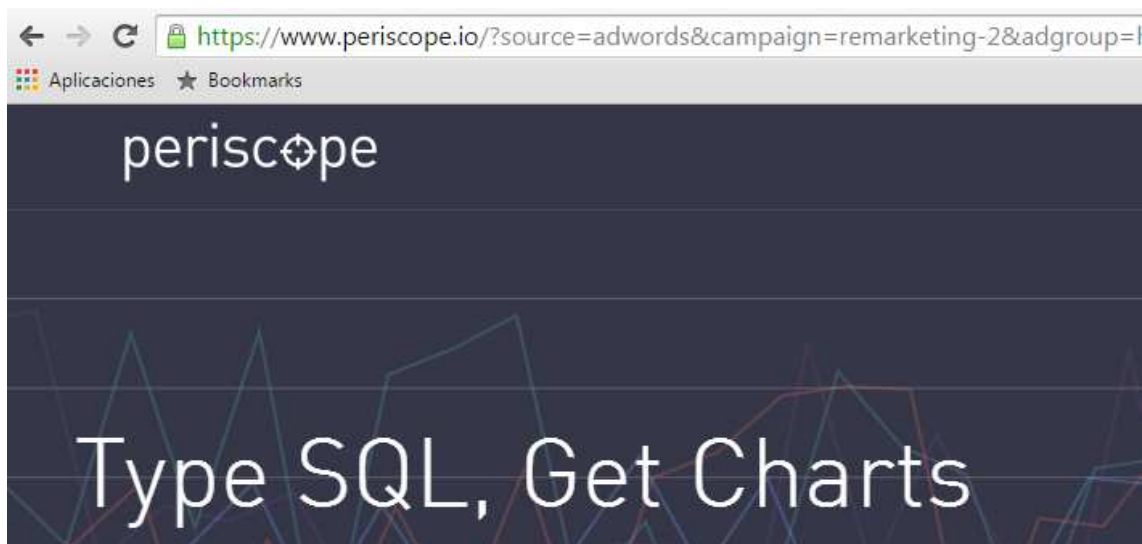
```
<a id="aw0" target="_blank" href="http://www.googleadservices.com/pagead/aclk?sa=L&ai=Cb0__ZrADVpaBLNi6Wpe2pZgP9-ε"
```

Como podéis observar el anunciante es periscope, sin embargo la URL que aparece como referencia es la del intermediario googleadservices.

La primera solución fue coger la URL y abrir una nueva pestaña con dicha dirección,



y obtener de esta manera la URL final.



El problema fundamental de esta solución es que estábamos obligando al usuario a navegar por una página que no había sido solicitada, con lo cual estaríamos condicionando el estudio. En nuestra base de datos aparecerían URLs que el usuario no había visitado, las cuales se estarían insertando como consecuencia del problema anterior.

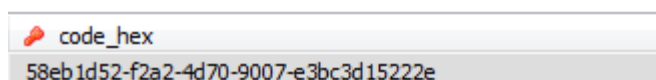
Además esta solución retrasaría la carga de la página por parte del navegador afectando a la experiencia de navegación del usuario, ya que por cada enlace publicitario encontrado en la página web implicaría una nueva petición URL.

La solución final fue la de realizar un proceso en el lado del servidor que cada cierto tiempo analizara las URLs almacenadas en la base de datos en busca de la URL final, de esta manera no condicionamos el estudio ya que no realizamos visitas a páginas innecesarias en el lado del cliente y mejoramos la experiencia de navegación evitando la carga en el navegador de cada enlace publicitario.

- **Pérdida del identificador del usuario**

Todas las páginas visitadas, así como sus cookies y sus enlaces publicitarios están asociadas a un usuario, cada vez que accedemos a una URL desde nuestro ordenador dicha página se almacenara para un usuario en concreto y dicho usuario no puede cambiar, debe ser único.

La primera solución que se toma para crear usuarios únicos fue la de crear un UUID (universal unique identifier), con esto nos asegurábamos de que no se repitiera el identificador del usuario



El problema surge cuando se cierra y se abre el navegador ya que se vuelve a generar un nuevo identificador, teniendo en este caso dos o más identificadores por ordenador cuando solo queremos que haya uno.

La solución final ha sido la de almacenar el identificador del usuario/equipo en un fichero (userKey.txt) en el lado del cliente, de esta forma cada vez que se cargue una página en el navegador se obtendrá el identificador leyendo dicho fichero. El fichero se rellena con el id del usuario cuando se instala el plugin en el navegador.

7. REFERENCIAS

7.1. Herramientas relacionadas con la publicidad

Lightbeam -> <https://addons.mozilla.org/en-us/Firefox/addon/lightbeam/>

Ghostery -> <https://www.ghostery.com/en/>

Adblockplus -> <https://adblockplus.org/>

Privacy Bagger -> <https://www.eff.org/privacybadger>

Privoxy -> <http://www.privoxy.org/>

Awazza -> <http://www.awazza.com/web/>

7.2. Bibliografía

- (1) Información utilizada para el capítulo 3.1.1 XAMMP

<https://es.wikipedia.org/wiki/XAMPP>

- (2) Información utilizada para el capítulo 3.1.2 APACHE

https://es.wikipedia.org/wiki/Servidor_HTTP_Apache

- (3) Información utilizada para el capítulo 3.1.3 MySQL

<https://es.wikipedia.org/wiki/MySQL>

- (4) Información utilizada para el capítulo 3.1.4 PHP

<https://es.wikipedia.org/wiki/PHP>

- (5) Información utilizada para el capítulo 3.1.5 ECLIPSE

[https://es.wikipedia.org/wiki/Eclipse_\(software\)](https://es.wikipedia.org/wiki/Eclipse_(software))

- (6) Información utilizada para el capítulo 3.1.6 JAVA

[https://es.wikipedia.org/wiki/Java_\(lenguaje_de_programación\)](https://es.wikipedia.org/wiki/Java_(lenguaje_de_programación))

(7) Información utilizada para el capítulo 3.2.1 FIREFOX

https://es.wikipedia.org/wiki/Mozilla_Firefox

(8) Información utilizada para el capítulo 3.2.2 AJAX

<https://librosweb.es/libro/ajax/>

(9) Información utilizada para el capítulo 3.2.3 JAVASCRIPT

<http://librosweb.es/libro/javascript/>

(10) Información utilizada para el capítulo 3.2.4 EXTENSIONES

<https://developer.mozilla.org/en-US/Add-ons>

7.3. Imágenes

Imagen utilizada en la introducción

<http://trafficbuilders.us/portfolio/ecommerce/>

Imagen utilizada en el capítulo de objetivos

<http://imagestudio47inked.com/2014/11/10/social-media-background-checks-precaution-or-invasion-of-privacy/>

Imágenes utilizadas en el capítulo de herramientas relacionadas

<https://www.mozilla.org/en-US/lightbeam/>

<https://www.ghostery.com/en/>

<https://adblockplus.org/>

<https://www.eff.org/privacybadger>

<http://www.privoxy.org/>

<http://www.awazza.com/web/>

Imágenes utilizadas en el diseño de la arquitectura funcional.

http://mib.isdi.es/sites/default/files/modulos/modulo3_0.png

https://www.mozilla.org/es-ES/Firefox/new/?utm_source=google&utm_medium=paidsearch&utm_campaign=sem2015Q3

<http://rjapproves.com/wp-content/uploads/2012/08/iconDB.jpg>

<http://www.monografias.com/trabajos89/poo-y-mvc-php/image005.png>

<http://www.invbit.com/wp-content/uploads/2010/09/nube.jpg>

http://www.kreamos.com/wp-content/uploads/2014/03/ICON_PROCESOS.png

ANEXO 1. MANUAL DEL DESARROLLADOR

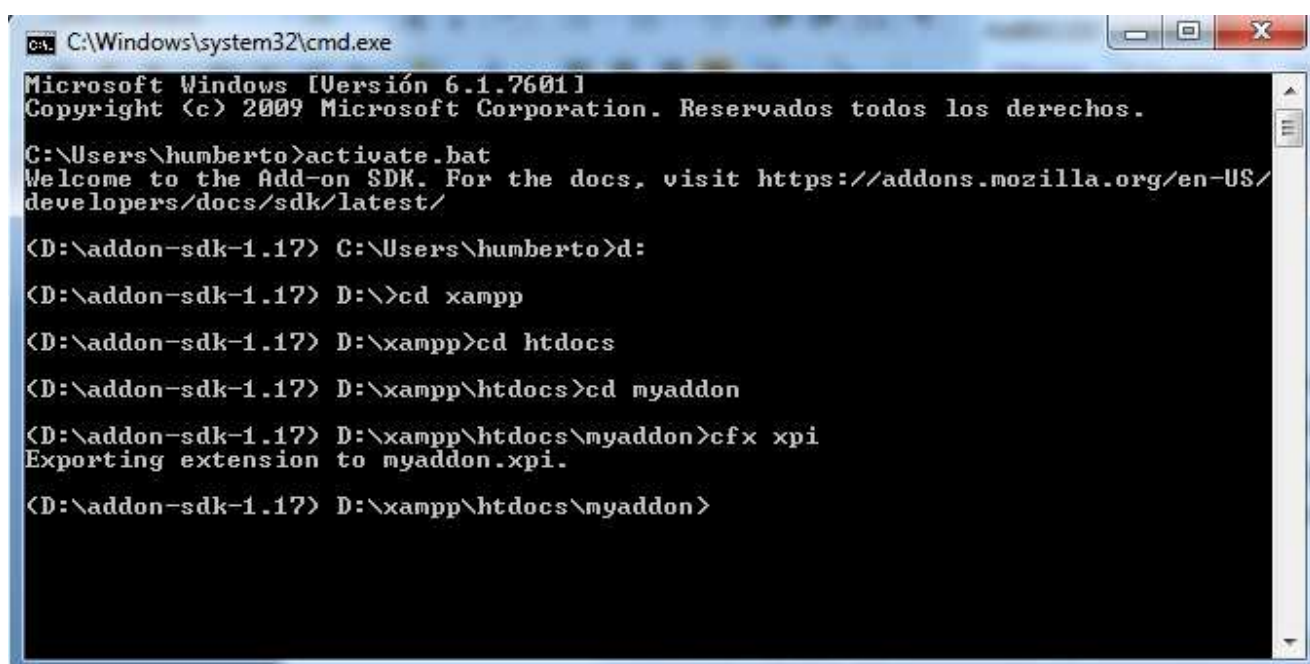
A continuación se detallaran los pasos a seguir por parte del desarrollador para la instalación tanto en el cliente como en el servidor del servicio que le permitirá recoger la información de navegación de los diferentes usuarios.

Paso 1: Configuración del plugin antes de la instalación en el navegador del cliente

El plugin requiere dos parámetros de configuración que son la URL del servidor de PHP y la ruta del fichero en el cual se va a generar el identificador de usuario. Estos parámetros se configuran modificando las dos primeras líneas del archivo main.js situado en la carpeta de myaddon.

```
//Ruta en la cual se almacenara el identificador del usuario  
var pathUserKey = "D:\\xampp\\htdocs\\myaddon\\userKey.txt";  
//URL del servidor de php que realizara las operaciones con la base de datos.  
var urlServer = "http://localhost/pfc/main.php"
```

Una vez que hayamos modificado las variables anteriores con los valores deseados tenemos que generar nuevamente el archivo.xpi que se instalara en el navegador, para generar el archivo debemos tener instalado el SDK Add-on (véase en el anexo correspondiente), con el SDK Add-on instalado abrimos un cmd y ejecutamos el comando activate.bat para activar la consola del sdk, posteriormente nos situamos en la carpeta que contiene el plugin (myaddon) y lanzamos el comando cfx xpi.



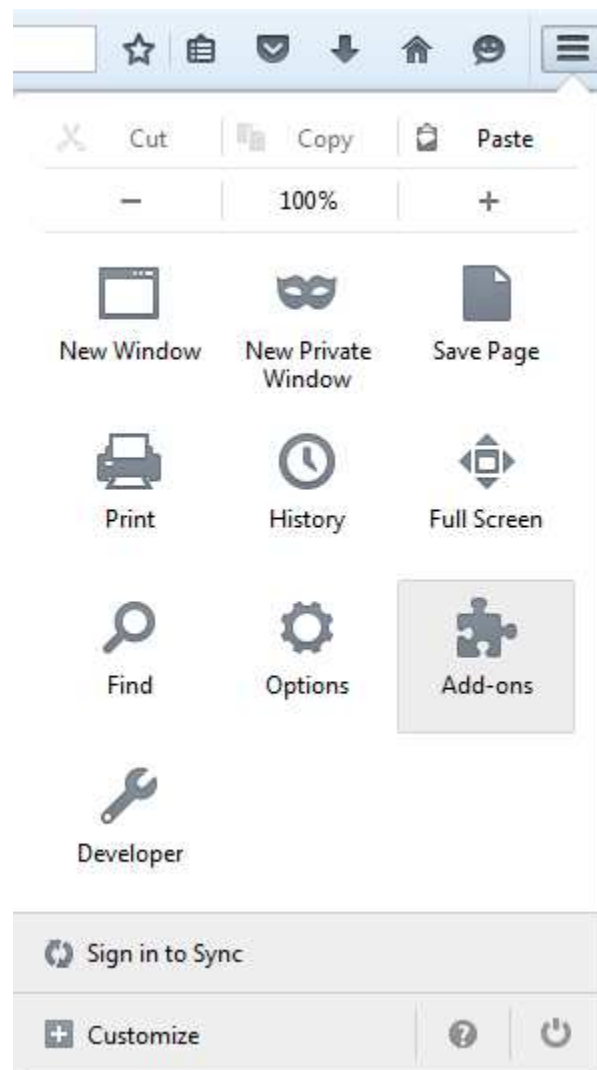
```
C:\Windows\system32\cmd.exe  
Microsoft Windows [Versión 6.1.7601]  
Copyright (c) 2009 Microsoft Corporation. Reservados todos los derechos.  
  
C:\Users\humberto>activate.bat  
Welcome to the Add-on SDK. For the docs, visit https://addons.mozilla.org/en-US/  
developers/docs/sdk/latest/  
  
<D:\addon-sdk-1.17> C:\Users\humberto>d:  
<D:\addon-sdk-1.17> D:\>cd xampp  
<D:\addon-sdk-1.17> D:\xampp>cd htdocs  
<D:\addon-sdk-1.17> D:\xampp\htdocs>cd myaddon  
<D:\addon-sdk-1.17> D:\xampp\htdocs\myaddon>cfx xpi  
Exporting extension to myaddon.xpi.  
<D:\addon-sdk-1.17> D:\xampp\htdocs\myaddon>
```


Paso 2: Instalación del plugin en el cliente

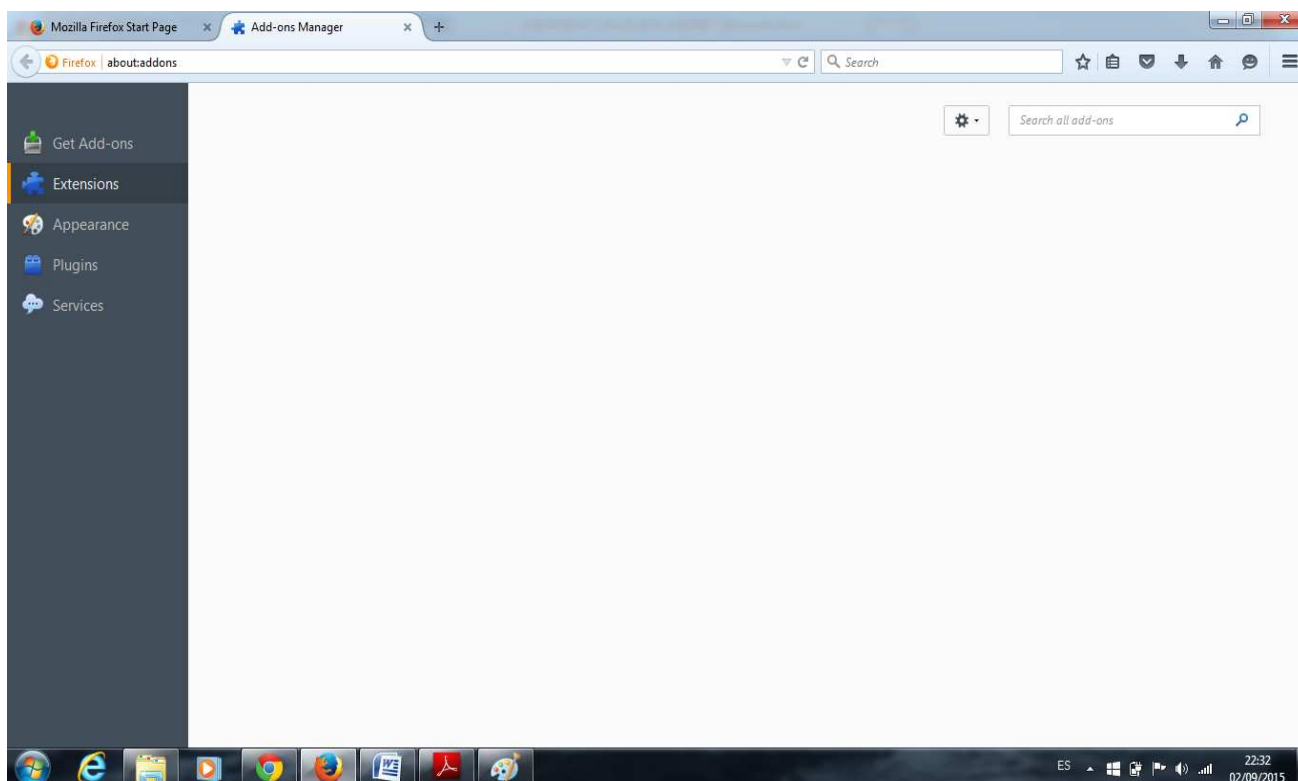
En el caso del cliente será necesario tener instalado el navegador Firefox, el cual se puede descargar a través del siguiente enlace:

<https://www.mozilla.org/en-US/Firefox/new/>

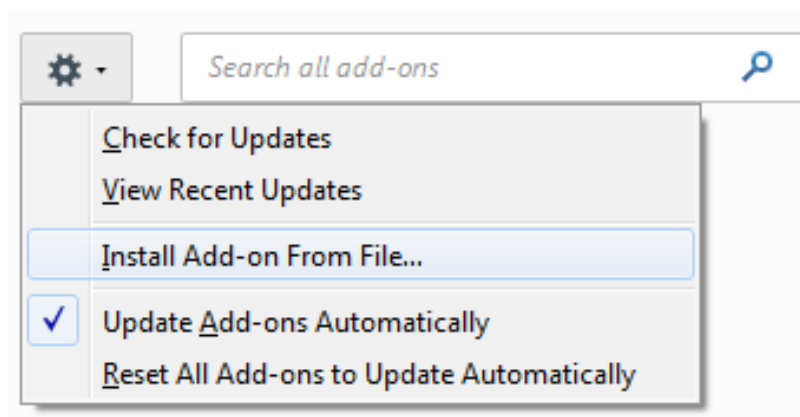
Una vez tengamos instalado el navegador en el equipo del cliente incluimos el plugin que recogerá la información, para su instalación abrimos el menú situado en la parte superior derecha del navegador.



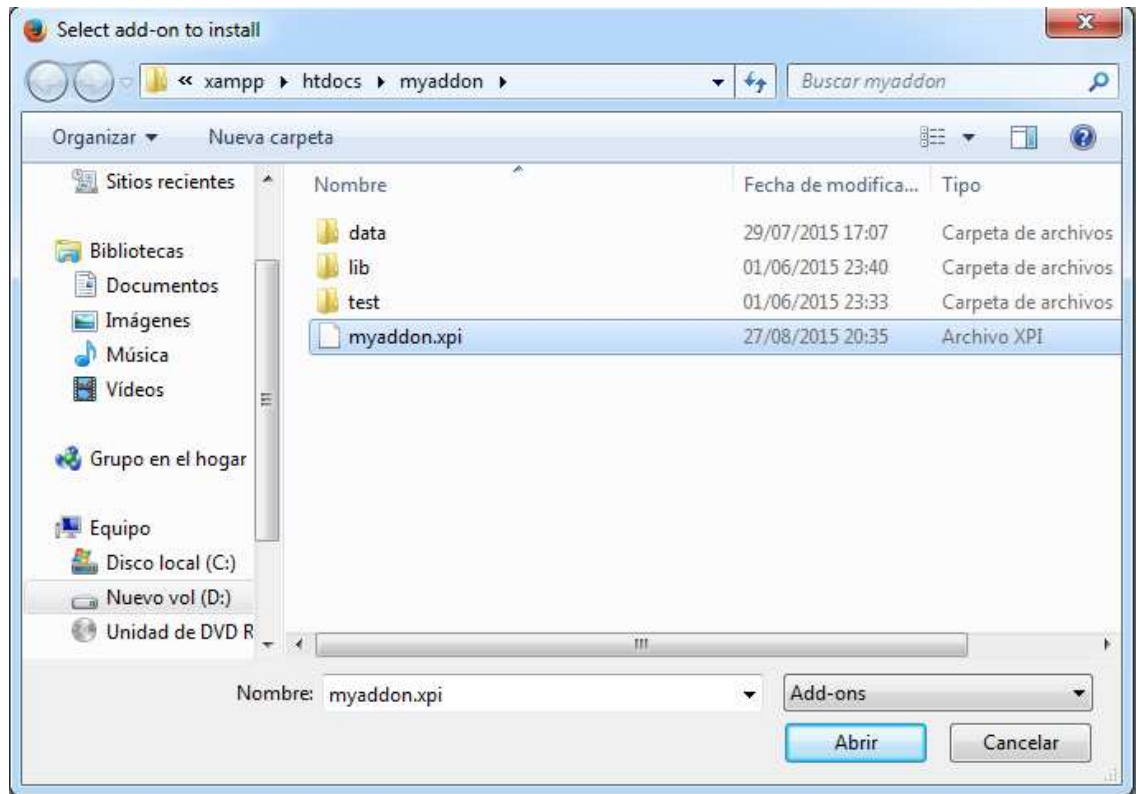
A continuación pinchamos en Add-ons y se nos abrirá la siguiente pestaña en el navegador.



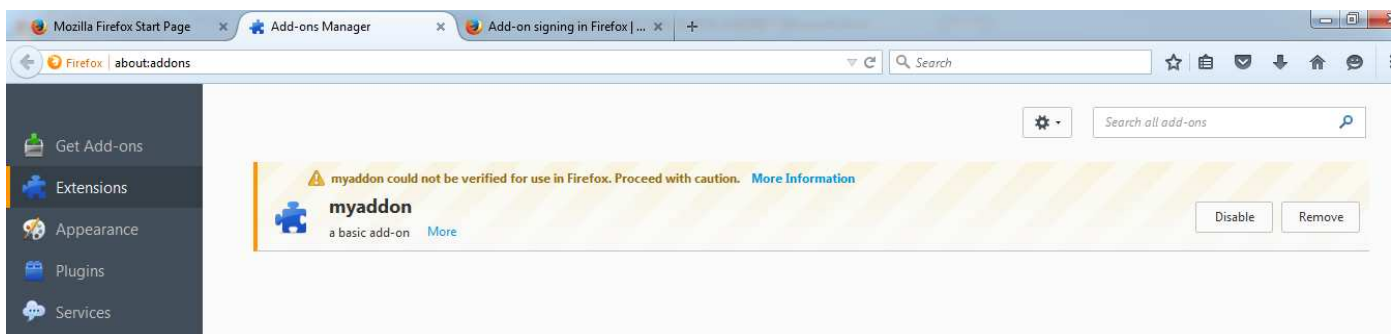
Como podéis ver aparecen seleccionadas las extensiones en el menú situado a la izquierda del navegador, para proceder a su instalación debemos hacer click en la rueda situada a la izquierda de la caja de texto de búsqueda, desplegándose el siguiente menú.



En el menú mostrado anteriormente seleccionaremos la opción de instalación, se nos abrirá el explorador de archivos para que elijamos el fichero.xpi que contiene el plugin a instalar.



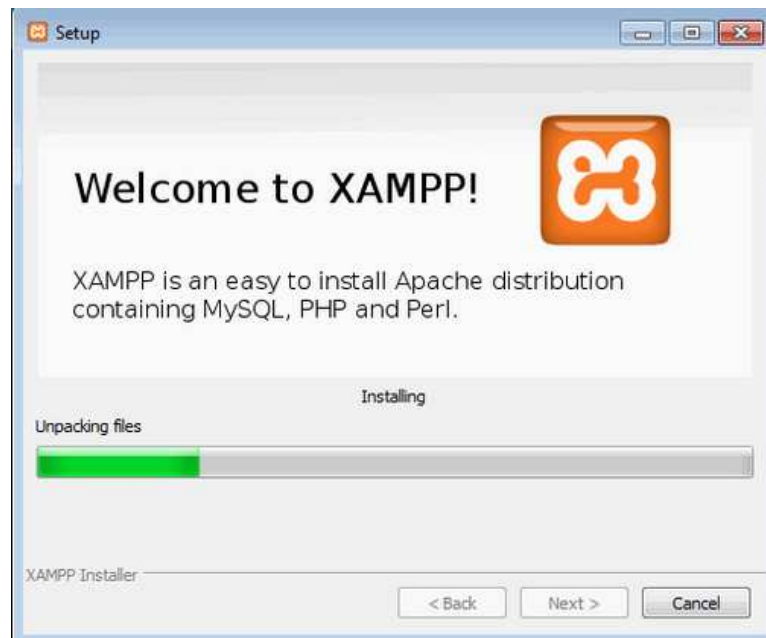
Una vez que seleccionemos el plugin en el explorador de archivos y le demos al botón de abrir se realizara la instalación en el navegador, si todo ha ido correctamente nos aparece en la pestaña del navegador el plugin instalado.



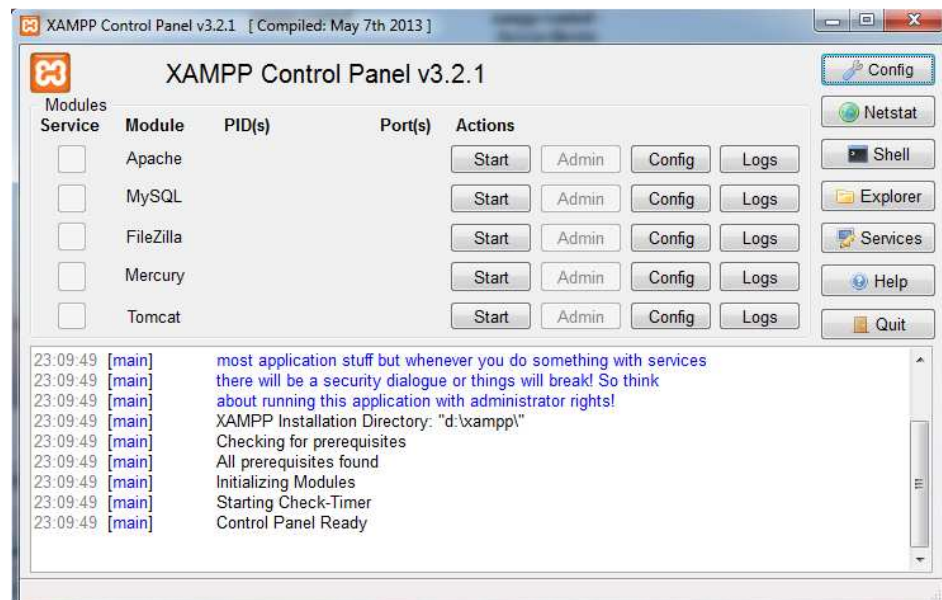
Esto es todo lo que necesitamos hacer en la parte del cliente para que el servicio que recoge la información del navegador funcione correctamente.

Paso 3: Instalación del Servidor

El servidor será el encargado de recopilar toda la información que le envía el cliente a través del plugin, para ello se instalara XAMPP que como ya hemos mencionado con anterioridad es un servidor que contiene un sistema de bases de datos MySQL, el servidor web Apache y un intérprete de PHP.

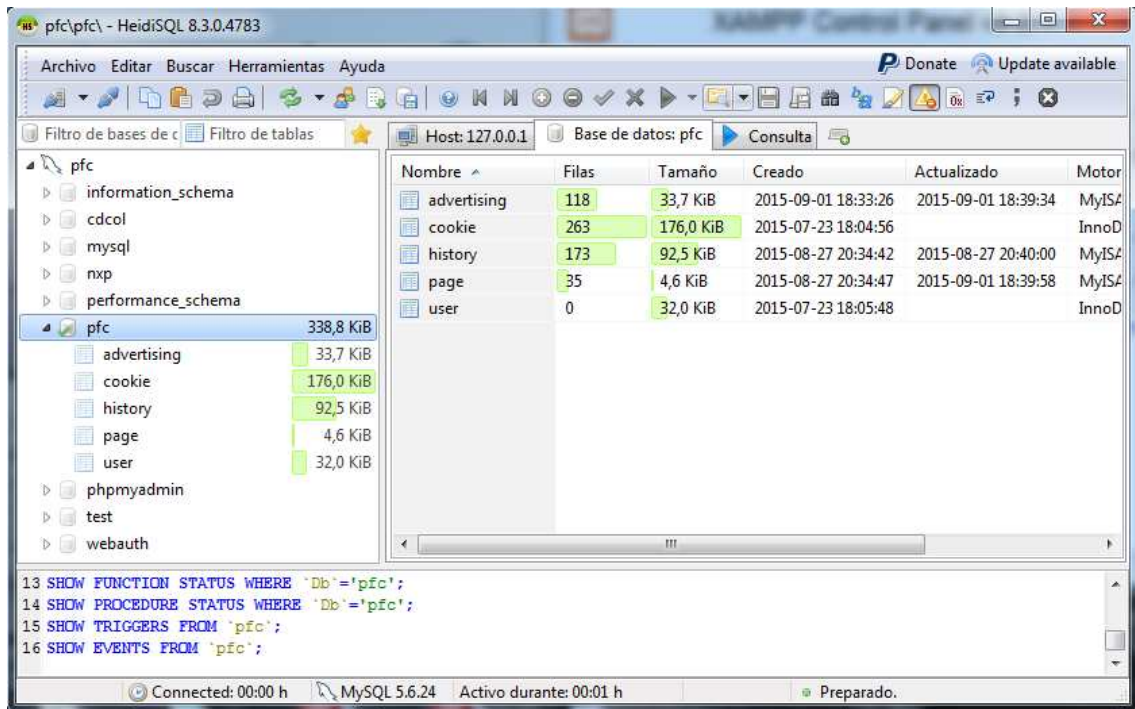


Una vez instalado XAMPP podemos arrancar los diferentes servicios a través de su panel de control que mostraremos a continuación.

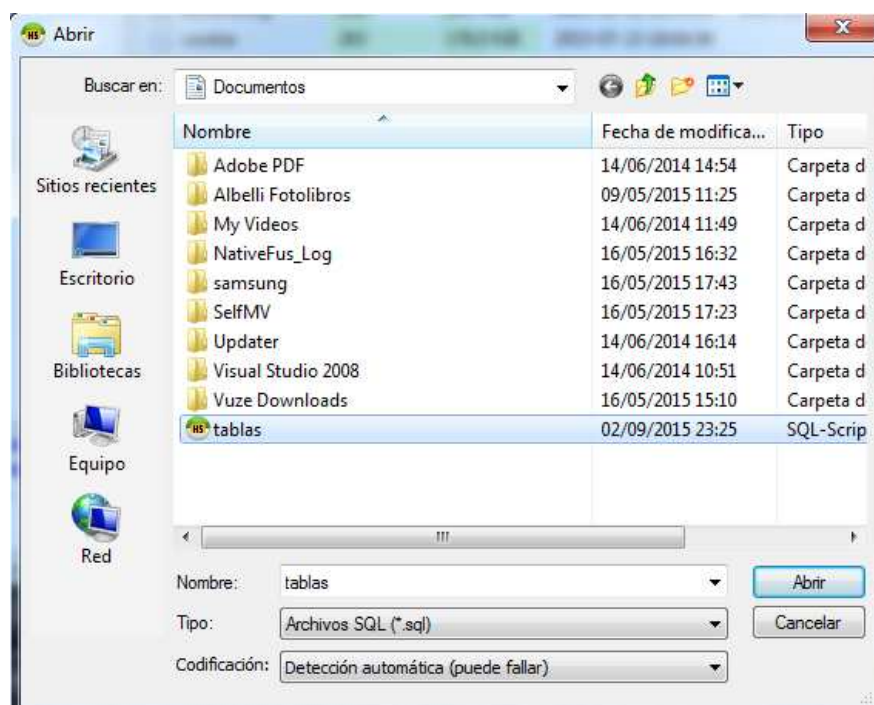


Para el proyecto en cuestión los servicios que vamos a necesitar arrancar son los correspondientes a servidor Apache y al sistema de gestión de base de datos MySQL.

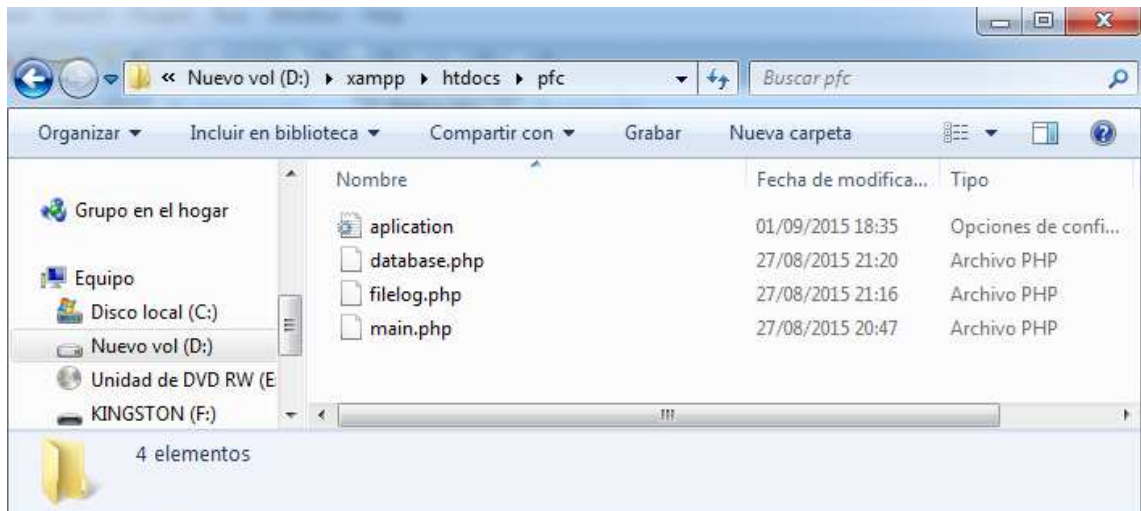
Una vez arrancado el servicio de MySQL procedemos a crear las diferentes tablas de la base de datos en las cuales se almacenaran toda la información de las trazas de navegación del usuario a través de la web, para la creación de dichas tablas se proporcionara un script MySQL el cual se puede ejecutar utilizando el comando MySQL a través de la consola o bien abriendo el HeidiSQL e importando el fichero como se muestra a continuación.



El fichero a importar se cargará haciendo click en el icono de explorador de archivos situado en el menú. Una vez abierta la ventana seleccionamos el fichero.sql que vamos a cargar.



Con las tablas creadas procederemos a incorporar en la carpeta htdocs los archivos.php que se encargaran de las operaciones de inserción en la base de datos de toda la información que nos envía el plugin referente a la navegación para un cliente determinado. Todos estos archivos se guardan en una carpeta denominada pfc la cual está dentro de htdocs.



El archivo aplicacion contiene una serie de parámetros de configuración que se detallaran a continuación.

[DATABASE]

HOST=localhost

USER=root

PASSWORD=1234

NAMEDB=pfc

[RunNow]

Frecuency=60

[LOG]

DIRECTORY_LOG_PHP=D:\xampp\htdocs\logs

DIRECTORY_LOG_JAVA=D:\\logjava\\logs

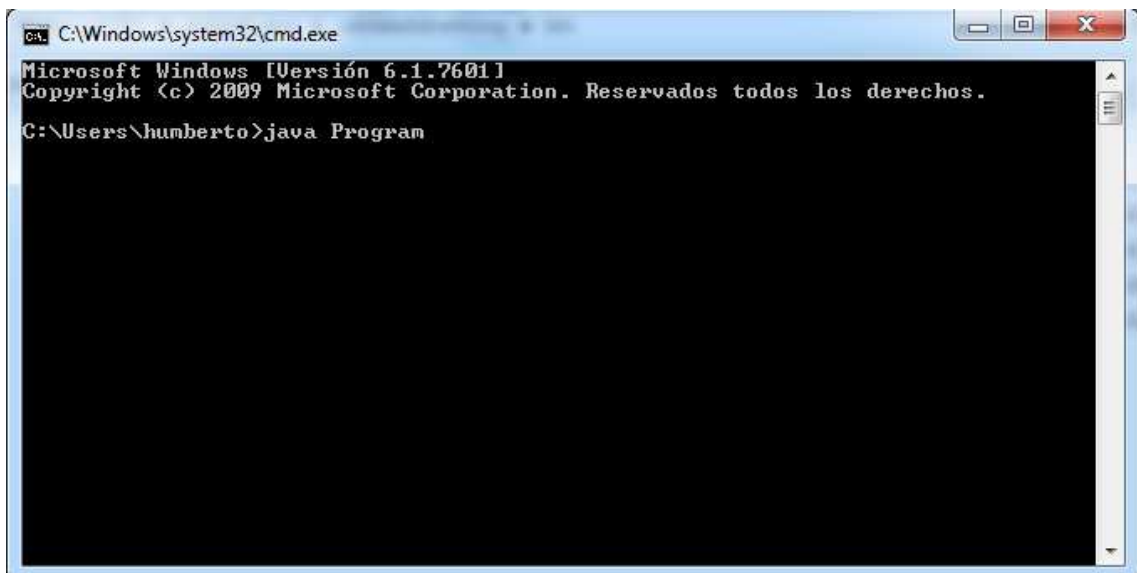
Como podéis ver hay tres partes bien diferenciadas (DATABASE, RunNow y LOG).

La primera configura los parámetros de la base de datos en la cual se va a insertar la información necesaria para el estudio.

En segundo lugar se configura la frecuencia con la cual se lanzara el servicio que nos permitirá obtener la URL final del anunciante.

En último lugar aparecen configurados las carpetas en las cuales se almacenaran las trazas de la aplicación, tanto para la parte de recogida de información realizada en PHP como para la parte de la obtención de la URL final realizada en Java, de la cual hablaremos a continuación.

Finalmente nos quedaría el servicio realizado en java validateAdvertising, el cual obtendrá la URL final, para lanzar el servicio abriremos una consola cmd.



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Versión 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Reservados todos los derechos.
C:\Users\humberto>java Program
```

Este servicio que obtiene la URL final se ejecutara cada x segundos, este tiempo es configurable y se puede modificar a través del archivo de configuración, concretamente en la clave denominada frequency.

ANEXO 2. INSTALACIÓN DEL SDK ADD-ON

A continuación se detallaran los pasos a seguir para la instalación del SDK add-on, el cual nos permitirá desarrollar plugins para Firefox.

Para desarrollar add-ons usando el SDK Add-on, necesitarás:

- Python 2.5, 2.6 o 2.7. Hay que tener en cuenta que las versiones 3.x de Python no son soportadas en ninguna plataforma. Asegúrate que Python se encuentre en la ruta del sistema. Para el desarrollo del proyecto se ha utilizado la versión 2.6
- Firefox.
- El propio SDK: puedes descargar la última versión estable del SDK como un tarball o como un archivo zip a través de la siguiente URL <https://developer.mozilla.org/es/docs/Mozilla/Add-ons/SDK/Tutorials/Installation>.

Instalación del SDK Add-on en Windows:

Extrae el contenido del archivo en el lugar de tu preferencia, y navega hasta el directorio principal de SDK con la terminal. Por ejemplo:

```
7z.exe x addon-sdk.zip
```

```
cd addon-sdk
```

Luego ejecuta:

```
bin\activate
```

En la línea de entrada de tu terminal ahora deberías tener un prefijo con el nombre completo del directorio raíz del SDK:

```
(C:\Users\mozilla\sdk\addon-sdk) C:\Users\Work\sdk\addon-sdk>
```

La orden activate establece algunas variables de entorno necesarias para el SDK solo para la terminal actual. Si abres un terminal nuevo, el SDK no estará activo hasta que ejecutes nuevamente activate. Si deseas una activación permanente vea el punto x.

Establecer activate permanentemente

Esto significa que puedes tener múltiples copias del SDK en diferentes partes del disco y cambiar entre ellas, o incluso tenerlas activas al mismo tiempo en diferentes terminales.

Establecer estas variables de forma permanente en tu entorno para que cada nuevo terminal pueda leerlas, podrás hacer que la activación sea permanente. Por lo tanto no tendrás que ejecutar activate cada vez que habrás una nueva terminal.

Debido a que el conjunto exacto de variables pueden cambiar en nuevas versiones del SDK, lo mejor es consultar la secuencia de órdenes de activación para determinar cuáles variables deben ser establecidas. La activación usa diferentes secuencias de órdenes de activación para entornos bash (Linux and OS X) y ambientes Windows. A continuación se presentara la activación para Windows, si desea configuración para otros entornos diríjase a la siguiente URL

<https://developer.mozilla.org/es/docs/Mozilla/Add-ons/SDK/Tutorials/Installation>

Windows

En Windows, bin\activate usa activate.bat, y puedes hacer la activación permanente usando la herramienta setx del terminal o el Panel de Control.

Pasos para el establecimiento permanente:

Paso 0. Abre el Panel de Control > Sistema > Avanzado > Variables de Entorno.

Paso 1. Agrega el directorio Python y el directorio SDK\bin a la Variable de la Ruta del Sistema.

Path=...;C:\PYTHON26\;C:\Users\Omid\Documents\FirefoxAddons\addon-sdk-1.15\bin

Paso 2. Agrega una Variable de usuario llamada CUDDLEFISH_ROOT y establece su valor a "C:\Users\Omid\Documents\FirefoxAddons\addon-sdk-1.15"

Paso 3. Agrega una Variable de usuario llamada PYTHONPATH y establece su valor a "C:\Users\Omid\Documents\FirefoxAddons\addon-sdk-1.15\python-lib"

Paso 4. Agrega una Variable de usuario llamada VIRTUAL_ENV y establece su valor a "C:\Users\Omid\Documents\FirefoxAddons\addon-sdk-1.15"

ANEXO 3. CREACIÓN DE UN ADD-ON

Inicializar un add-on vacío

En la terminal, cree un directorio nuevo. El directorio no tiene que estar bajo la raíz del SDK: puede crearlo donde quiera. Diríjase al directorio recién creado, escriba `cfx init`, y presione la tecla enter o intro:

```
mkdir my-addon
```

```
cd my-addon
```

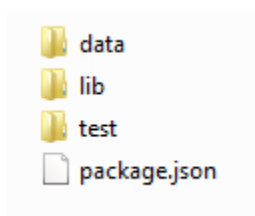
```
cfx init
```

Vea una salida como esta:

```
* lib directory created
* data directory created
* test directory created
* generated jid automatically: jid1-KWqN3N4HE2itNQ
* package.json written
* test/test-main.js written
* lib/main.js written
```

Your sample add-on is now ready.

Do "cfx test" to test it and "cfx run" to try it. Have fun!



Esto creará una serie de directorios necesarios para la creación del Add-on. A continuación ya puede empezar a codificar el add-on para ello deberá escribir el código necesario en el archivo `main.js` situado en el directorio `lib`.

En el caso de que sea necesario mostrar imágenes en el add-on deberá incluirlas en el directorio `data`.

Una vez que tenga construido el add-on vuelva al terminal y escriba lo siguiente

```
cfx run
```

Esta es la orden del SDK para ejecutar una nueva instancia de Firefox con el add-on instalado, en el terminal donde se ejecuta dicha orden aparecerán los posibles errores en la codificación del add-on en el caso de que los haya.

Empaquetar el add-on

Cuando el add-on este terminado y listo para distribuirse, será necesario empaquetarlo como un archivo XPI. Este es el formato de archivo de instalación para los add-ons de Firefox. Puede distribuir los archivos XPI por cuenta propia o publicarlos en <https://addons.mozilla.org> para que otros usuarios puedan descargarlos e instalarlos.

Para construir un XPI, solo se debe ejecutar la orden `cfx xpi` desde el directorio del add-on:

```
cfx xpi
```

Vera una salida como esta:

```
Exporting extension to my-addon.xpi.
```

Para comprobar que funcione, intente instalar el archivo XPI en el Firefox instalado en su sistema. Puede hacer esto presionando la combinación de teclas Ctrl+O (Cmd+O en Mac) desde el Firefox, o seleccionando la opción "Abrir" desde el menú "Archivo" de Firefox. Esto abrirá un dialogo de selección; navegue hasta el archivo "my-addon.xpi", abra el archivo y siga las instrucciones para instalar el add-on.

Resumen

En este tutorial se ha construido y empaquetado un add-on usando tres órdenes:

- `cfx init` para inicializar una plantilla vacía para el add-on.
- `cfx run` para ejecutar una instancia nueva de Firefox con el add-in instalado, para poder realizar pruebas.
- `cfx xpi` para empaquetar el add-on en un archivo XPI para su distribución.

Estas son las tres órdenes principales usadas cuando se desarrollan add-ons con SDK.

ANEXO 4. LIBRERÍAS UTILIZADAS EN EL DESARROLLO DEL ADD-ON

SDK Add-on nos proporciona una serie de librerías javascript que nos permitirán construir nuestro add-on con las funcionalidades necesarias. Las librerías del SDK se dividen en librerías de alto nivel y librerías de bajo nivel. A continuación se nombrarán las librerías utilizadas en la realización del proyecto.

TABS

Tabs es una de la librerías de alto nivel que nos proporciona el SDK Add-on, esta librería nos ofrece funciones que nos permiten abrir nuevas pestañas en el navegador y manipular su contenido, además incluye una serie de eventos que nos notifican cuando una pestaña se cierra, se activa, se desactiva o cuando el contenido de la página que hemos abierto en dicha pestaña ha terminado de cargarse. Este ultimo evento ha sido el que hemos utilizado en el proyecto, ya que nos permite saber cuando un usuario ha cargado una página web en el navegador.

IO/TEXT-STREAMS

Esta librería del sdk nos proporciona dos funciones, una para la lectura de ficheros y otra para la escritura (TextReader y TextWriter). Estas son las funciones que se utilizan para guardar el identificador del usuario, cuando se instala el plugin se genera un identificador único de usuario, el cual se almacena en el fichero que aparece en la ruta especificada en la variable pathUserKey del archivo main.js. Este identificador es el que se utilizara para enviar la información de las páginas visitadas al servidor que posteriormente las almacenará en la base de datos, cada vez que el usuario visita una página debemos obtener su identificador leyendo el fichero.

PLACES

Places es un sistema de gestión de marcadores e historial introducido en firefox 3, este sistema nos ofrece una mayor flexibilidad y nos permite realizar consultas complejas sobre los sitios que ha visitado el usuario. También incluye nuevas características, incluyendo el almacenamiento favicon (**imagen de 16x16 o 32x32 píxeles** que sirve de reconocimiento de tu página web) y la capacidad para anotar páginas con información arbitraria, además introduce nuevos interfaces de usuario para la gestión de toda esta información.

nsINavHistoryService es el query object que utiliza Places para consultar información acerca del historial del usuario.

El primer paso es crear la consulta con sus opciones , dichas opciones se rellenan con los parámetros que deseas . Utilice nsINavHistoryService.getNewQuery () y nsINavHistoryService.getNewQueryOptions () para recuperar los objetos vacíos. Los valores predeterminados para estos objetos se traducirá en una consulta que devuelve todo su historial de navegación en una lista plana, en el caso de nuestro proyecto es justamente lo que deseamos obtener. La consulta se realiza llamando a la función historyService.executeQuery(query, options).

El resultado de la consulta nos devuelve un array de objetos `nsINavHistoryResultNode`, dicho objeto posee una serie de atributos con información acerca de la página visitada, se pueden ver todos y cada uno de los atributos visitando la página.

<https://developer.mozilla.org/en-US/docs/Mozilla/Tech/XPCOM/Reference/Interface/nsINavHistoryResultNode>

En el caso del proyecto se han utilizado dos propiedades:

- **URI:** En esta propiedad se almacena la URL de la página.
- **TIME:** Almacena la fecha y la hora en la cual el usuario ha accedido a la página web, esta propiedad es solo de lectura

CONTENT SCRIPT

La mayoría de los Add-ons necesitan acceder y modificar el contenido de las páginas web, el código del add-on alojado en el archivo `main.js` no tiene acceso directo a los elementos de la página. El SDK Add-on separa el código que gestiona el acceso al contenido web en scripts denominados `content-script`. A continuación describiremos como desarrollar e implementar `content-script`.

Los `content-script` pueden ser uno de los aspectos más confusos a la hora de trabajar con el SDK Add-on, para su utilización deberá tener en cuenta los siguientes principios básicos.

- El código principal de Add-on incluido en el `main.js` y en otros módulos (carpeta `lib`) solo pueden utilizar las librerías de alto nivel y las de bajo nivel, no tienen acceso al contenido de la página.
- Los `content-script` no pueden acceder al SDK API, sin embargo si tienen acceso al contenido de la página.
- Las librerías del SDK API que utilizan `content-script` como por ejemplo `tabs` y `page-mod` ofrecen funciones que nos permiten acceder al contenido de las páginas.
- Los `content-script` se pueden cargar incluyendo el código directamente o haciendo referencia a un fichero. A continuación se mostrará un ejemplo de los dos casos
 - Incluyendo el código directamente:

```

1 // main.js
2
3 var pageMod = require("sdk/page-mod");
4 var contentScriptValue = 'document.body.innerHTML = ' +
5                           ' "<h1>Page matches ruleset</h1>";';
6
7 pageMod.PageMod({
8   include: "*.mozilla.org",
9   contentScript: contentScriptValue
10 });

```

- A través de un fichero:

```

1 // main.js
2
3 var data = require("sdk/self").data;
4 var pageMod = require("sdk/page-mod");
5
6 pageMod.PageMod({
7   include: "*.mozilla.org",
8   contentScriptFile: data.url("content-script.js")
9 });

```

```

1 // content-script.js
2
3 document.body.innerHTML = "<h1>Page matches ruleset</h1>";

```

Un API de paso de mensajes permite que el código main y el content-script se comuniquen el uno con el otro.

A continuación se mostrarán las librerías de alto nivel del SDK Add-on que utilizan content-script para modificar las páginas web.

- Page-mod: Permite ejecutar content-script para todas aquellas URL que cumplan con un determinado patrón.
- Tabs: Se utiliza para trabajar con las pestañas del navegador, esta librería utiliza la función attach para cargar el content-script que nos permitirá acceder al contenido de la página.

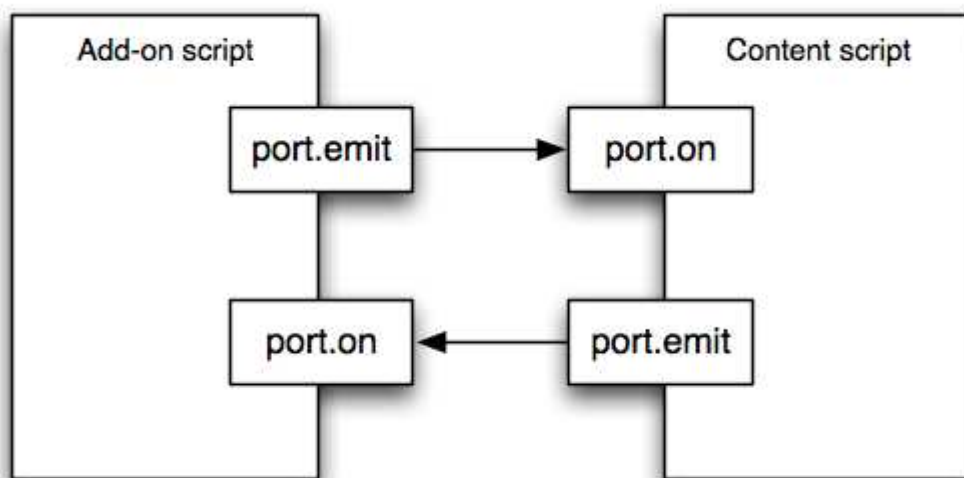
- Page-woker: Esta librería nos permite abrir una pestaña en segundo plano sin que esta se muestre al usuario, el content-script se utiliza para acceder y manipular los elementos de la página a través del DOM.
- Context-menu: Usa el content-script para interactuar con la página en la cual el menú ha sido invocado.

El `contentScriptWhen` es la opción que especifica cuando se realiza la llamada al content-script. Existe tres tipos de opciones de llamada:

- Start: Realiza la llamada al script una vez que el elemnto es insertado en el DOM de la página.
- Ready: Realiza la llamada al script una vez que el DOM de la página ha sido cargado.
- End: Realiza la llamada al script una vez que se ha cargado el DOM, JS, CSS e imágenes de la página.

Nota: El `tab.attach` no acepta el `contentScriptWhen`, ya que esta función se llama generalmente después de que la página ha sido cargada.

El intercambio de información del add-on con el content-script y viceversa se realiza a través del objeto `port`. El envío de mensajes del add-on al content-script se realiza a través de la función `port.emit()`, y para la recepción de los mensajes enviados por el content-script se emplea la función `port.on()`.



La comunicación de un content-script con otro es posible si ambos están incluidos en el mismo contexto. Por ejemplo si una sola llamada a `Tab.attach` usted incluye varios content-script estos podrán comunicarse entre ellos, pero si usted llama a `Tab.attach` dos veces adjuntando un content-script cada vez, en este caso los content-script no podrán comunicarse entre sí. A continuación se muestra una llamada en la cual se incluyen dos content-script (`operation.js` y `page.js`) en el mismo contexto.

`contentScriptFile: [self.data.URL("operations.js"), self.data.URL("page.js")]`